

FIGURE 1

CCAGGTCCAACCTGCACCTCGGTTCTATCGATTGAATTCCCCGGGGATCCTCTAGAGATCCCT
CGACCTCGACCCACGCGTCCGCCAAGCTGGCCCTGCACGGCTGCAAGGGAGGCTCCTGTGGA
CAGGCCAGGCAGGTGGGCCTCAGGAGGTGCCTCCAGGCGGCCAGTGGGCCTGAGGCCCCAGC
AAGGGCTAGGGTCCATCTCCAGTCCCAGGACACAGCAGCGGCCACCATGGCCACGCCTGGGC
TCCAGCAGCATCAGCAGCCCCCAGGACCGGGGGAGGCACAGGTGGCCCCCACCACCCGGAGG
AGCAGCTCCTGCCCCTGTCCGGGGGATGACTGATTCTCCTCCGCCAGGCCACCCAGAGGAGA
AGGCCACCCCGCCTGGAGGCACAGGCCATGAGGGGCTCTCAGGAGGTGCTGCTGATGTGGCT
TCTGGTGTGTCAGTGGGCGGCACAGAGCACGCCTACCGGCCCGGCCGTTAGGGTGTGTGCT
GTCCCGGGCTCACGGGGACCTGTCTCCGAGTCGTTTCGTGCAGCGTGTGTACCAGCCCTTCC
TCACCACCTGCGACGGGCACCGGGCCTGCAGCACCTACCGAACCATTATAGGACCGCCTAC
CGCCGCAGCCCTGGGCTGGCCCCCTGCCAGGCCTCGCTACGCGTGCTGCCCCGGCTGGAAGAG
GACCAGCGGGCTTCCTGGGGCCTGTGGAGCAGCAATATGCCAGCCGCCATGCCGGAACGGAG
GGAGCTGTGTCCAGCCTGGCCGCTGCCGCTGCCCTGCAGGATGGCGGGGTGACACTTGCCAG
TCAGATGTGGATGAATGCAGTGCTAGGAGGGGGCGGCTGTCCCCAGCGCTGCATCAACACCGC
CGGCAGTTACTGGTGCCAGTGTGTTGGGAGGGGCACAGCCTGTCTGCAGACGGTACACTCTGTG
TGCCCAAGGGAGGGCCCCCAGGGTGGCCCCCAACCCGACAGGAGTGGACAGTGCAATGAAG
GAAGAAGTGCAGAGGCTGCAGTCCAGGGTGGACCTGCTGGAGGAGAAGCTGCAGCTGGTGCT
GGCCCCACTGCACAGCCTGGCCTCGCAGGCACTGGAGCATGGGCTCCCGGACCCCGGCAGCC
TCCTGGTGCACCTCCTTCCAGCAGCTCGGCCGCATCGACTCCCTGAGCGAGCAGATTTCTTTC
CTGGAGGAGCAGCTGGGGTCTGCTCCTGCAAGAAAGACTCGTGACTGCCCAGCGCCCCAGG
CTGGACTGAGCCCCTCACGCCGCCCTGCAGCCCCCATGCCCTGCCCAACATGCTGGGGGTC
CAGAAGCCACCTCGGGGTGACTGAGCGGAAGGCCAGGCAGGGCCTTCCTCCTTTTCTCTCCTC
CCCTTCCCTCGGGAGGGTCCCCAGACCCTGGCATGGGATGGGCTGGGATTTTTTTTTGTGAAT
CCACCCCTGGCTACCCCCACCCTGGTTACCCCAACGGCATCCCAAGGCCAGGTGGGCCCTCA
GCTGAGGGAAGGTACGAGTTCCCCTGCTGGAGCCTGGGACCCATGGCACAGGCCAGGCAGCC
CGGAGGCTGGGTGGGGCCTCAGTGGGGGCTGCTGCCTGACCCCCAGCAATAAAAATGAAA
CGTGAAAGGGCGGCCGCGACTCT
AGAGTCGACCTGCAGAAGCTTGGCCGCCATGGCCCAACTTGTTTATTGCAGCTTATAATGGT
TACAAAT

FIGURE 2

MTDSPPPGHPPEEKATPPGGTGHEGLSGGAADVASGVGSGRHRARLPARPLGCVLSRAHGDPV
SESFVQRVYQPFLTTCDGHRACSTYRTIYRTAYRRSPGLAPARPRYACCPGWKRTSGLPGAC
GAAICQPPCRNGGSCVQPGRCRCPAGWRGDTQCQSDVDECSARRGGCPQRCINTAGSYWCQCW
EGHSLSADGTLCVPKGGPPRVAPNPTGVDSAMKEEVQRLQSRVDLLEEKQLQLVLAPLHSLAS
QALEHGLPDPGSLLVHSFQQLGRIDSLSEQISFLEEQLGSCSCKKDS

Signal sequence:

amino acids 1-19

cAMP- and cGMP-dependent protein kinase phosphorylation sites.

amino acids 93-97, 270-274

N-myristoylation sites.

amino acids 19-25, 78-84, 97-103, 100-106, 103-109, 157-163,
191-197, 265-271

Amidation site.

amino acids 26-30

Aspartic acid and asparagine hydroxylation site.

amino acids 152-164

Cell attachment sequence.

amino acids 130-133

EGF-like domain cysteine pattern signature.

amino acids 123-135

FIGURE 3

CGCTCGCCCCGTCGCCCCCTCGCCTCCCCGCAGAGTCCCCTCGCGGCAGCAGATGTGTGTGGG
GTCAGCCCACGGCGGGGACTATGGTGAAATTCGCGGCGCTCACGCACTACTGGCCCCCTGATC
CGGTTCTTGGTGCCCCCTGGGCATCACCAACATAGCCATCGACTTCGGGGAGCAGGCCTTGAA
CCGGGGCATTGCTGCTGTCAAGGAGGATGCAGTCGAGATGCTGGCCAGCTACGGGCTGGCGT
ACTCCCTCATGAAGTTCTTCACGGGTCCCATGAGTGACTTCAAAAATGTGGGCCTGGTGTTT
GTGAACAGCAAGAGAGACAGGACCAAAGCCGTCTGTGTATGGTGGTGGCAGGGGCCATCGC
TGCCGTCTTTCACACACTGATAGCTTATAGTGATTTAGGATACTACATTATCAATAAACTGC
ACCATGTGGACGAGTCGGTGGGGAGCAAGACGAGAAGGGCCTTCCTGTACCTCGCCGCCTTT
CCTTTCATGGACGCAATGGCATGGACCCATGCTGGCATTCTCTTAAACACAAATACAGTTT
CCTGGTGGGATGTGCCTCAATCTCAGATGTCATAGCTCAGGTTGTTTTTGTAGCCATTTTGC
TTCACAGTCACCTGGAATGCCGGGAGCCCCCTGCTCATCCCGATCCTCTCCTTGTACATGGGC
GCACTTGTGCGCTGCACCACCTGTGCCTGGGCTACTACAAGAACATTCACGACATCATCCC
TGACAGAAGTGGCCCCGAGCTGGGGGGAGATGCAACAATAAGAAAGATGCTGAGCTTCTGGT
GGCCTTTGGCTCTAATTCTGGCCACACAGAGAATCAGTCGGCCTATTGTCAACCTCTTTGTT
TCCCGGGACCTTGGTGGCAGTTCTGCAGCCACAGAGGCAGTGGCGATTTTGACAGCCACATA
CCCTGTGGGTACATGCCATACGGCTGGTTGACGGAAATCCGTGCTGTGTATCCTGCTTTTCG
ACAAGAATAACCCAGCAACAACTGGTGAGCACGAGCAACACAGTCACGGCAGCCACATC
AAGAAGTTCACCTTCGTCTGCATGGCTCTGTCACTCACGCTCTGTTTCGTGATGTTTTGGAC
ACCCAACGTGTCTGAGAAAATCTTGATAGACATCATCGGAGTGGACTTTGCCTTTGCAGAAC
TCTGTGTTGTTCCCTTTCGGATCTTCTCCTTCTTCCCAGTTCAGTCACAGTGAGGGCGCAT
CTCACCGGGTGGCTGATGACACTGAAGAAAACCTTCGTCTTGGCCCCAGCTCTGTGCTGCG
GATCATCGTCCTCATCGCCAGCCTCGTGGTCTACCCTACCTGGGGGTGCACGGTGGGACCC
TGGGCGTGGGCTCCCTCCTGGCGGGCTTTGTGGGAGAATCCACCATGGTGCATCGCTGCG
TGCTATGTCTACCGGAAGCAGAAAAGAAGATGGAGAATGAGTCGGCCACGGAGGGGGAAGA
CTCTGCCATGACAGACATGCCTCCGACAGAGGAGGTGACAGACATCGTGGAAATGAGAGAGG
AGAATGAATAAGGCACGGGACGCCATGGGCACTGCAGGGACGGTCAGTCAGGATGACACTTC
GGCATCATCTCTTCCCTCTCCCATCGTATTTTGTTCCTTTTTTTTGTGTTTGGTAAT
GAAAGAGGCCTTGATTTAAAGGTTTCGTGTCAATTCTCTAGCATACTGGGTATGCTCACACT
GACGGGGGGACCTAGTGAATGGTCTTTACTGTTGCTATGTAAAAACAAACGAAACAACTGAC
TTCATACCCCTGCCTCACGAAAACCCAAAAGACACAGCTGCCTCACGGTTGACGTTGTGTCC
TCCTCCCTGGACAATCTCCTCTTGGAACCAAAGGACTGCAGCTGTGCCATCGCGCCTCGGT
CACCTGACAGCAGGCCACAGACTCTCCTGTCCCCCTTCATCGCTCTTAAGAATCAACAGG
TTAAAACTCGGCTTCCTTTGATTTGCTTCCCAGTCACATGGCCGTACAAAGAGATGGAGCCC
CGGTGGCCTCTTAAATTTCCCTTCTGCCACGGAGTTCGAAACCATCTACTCCACACATGCAG
GAGGCGGGTGGCACGCTGCAGCCCGGAGTCCCCGTTCACTGAGGAACGGAGACCTGTGAC
CACAGCAGGCTGACAGATGGACAGAATCTCCCGTAGAAAGGTTTGGTTTGAAATGCCCCGGG
GGCAGCAAACCTGACATGGTTGAATGATAGCATTTCACTCTGCGTTCTCCTAGATCTGAGCAA
GCTGTCAGTTCTCACCCCCACCGTGTATATACATGAGCTAACTTTTTTAAATTGTACAAAA
GCGCATCTCCAGATTCCAGACCTGCCGCATGACTTTTCCTGAAGGCTTGCTTTTCCCTCGC
CTTTCCTGAAGGTGCGATTAGAGCGAGTCACATGGAGCATCCTAACTTTGCATTTTAGTTTT
TACAGTGAACCTGAAGCTTTAAGTCTCATCCAGCATCTAATGCCAGGTGCTGTAGGGTAAC
TTTTGAAGTAGATATATTACCTGGTTCTGCTATCCTTAGTCATAACTCTGCGGTACAGGTAA
TTGAGAATGTACTACGGTACTTCCCTCCCACACCATACGATAAAGCAAGACATTTTATAACG
ATACCAGAGTCACTATGTGGTCTCCCTGAAATAACGCATTCGAAATCCATGCAGTGCAGTA
TATTTTTCTAAGTTTTTGAAAGCAGGTTTTTCTTTAAAAAAATTATAGACACGGTTCCT
AAATTGATTTAGTCAGAAATTCCTAGACTGAAAGAACCTAAACAAAAAAATATTTTAAAGATA
TAAATATATGCTGTATATGTTATGTAATTTATTTTAGGCTATAATACATTTCTATTTTCGC
ATTTTCAATAAAATGTCTCTAATACAAAAA

FIGURE 4

MVKFPALTHYWPLIRFLVPLGITNIAIDFGEQALNRGIAAVKEDAVEMLASYGLAYSLMKFF
TGPMSDFKNVGLVFNNSKRDRTKAVLCMVVAGAIAAVFHTLIAYSDLGYIINKLHHVDES
GSKTRRAFLYLAAFPFMDAMAWTHAGILLKHKYSFLVGCASISDVIAQVVFVAILLHSHLEC
REPLIPILSLYMGALVRCTTLCLGYKNIHDIIPDRSGPELGGDATIRKMLSFWWPLALIL
ATQRISRPIVNLFVSRDLGGSSAATEAVAILTATYPVGHMPYGWLTEIRAVYPAFDKNNPSN
KLVSTSNVTAAHIKKFTFVCMALSLTLCFVMFWTPNVSEKILIDIIGVDFAFELCVVPLR
IFSFPVPVTVRAHLTGWLMTLKKTFVLAPSSVLRIIVLIASLVVLPYLGVHGATLGVGSLL
AGFVGESTMVAIAACYVYRKQKKMENESATEGEDSAMTDMPTTEEVTDIVEMRENE

Transmembrane domains:

amino acids 86-106, 163-179, 191-205, 237-253, 327-343, 357-374,
408-423, 431-445

[illegible]

CCTGACAGAAGTGTCCCGGAGCTGGGGGAGATNCAACATTAAGAAGATGCTGAGCTTCTGGT
GCCNTTTGGCTCTAATTCTGGCCACACAGAGAANCAGTCGGCCTATTGTCAACCTCTTTGTT
TCCCGGGACCTTGGTGGCAGTTCTGCAGCCACAGAGGCAGTGGCGATTTTGACAGCCACATA
CCCTGTGGGTACATGCCATACGGCTGGTTGACGGAAATCCGTGCTGTGTATCCTGCTTTTCG
ACAAGAATAACCCAGCAACAACTGGTGAGCACGAGCAACACAGTCACGGCGGCCACATC
AAGAAGTTCACCTTCGTCTGCATGGCTCTGTCACTCACGCTCTGTTTCGTGATGTTTTGGAC
ACCCAACGTGTCTGNAAAATCTTGATAGACATCATCGGAGTGGACTTTGCCTTTGCAGAAC
TCTGTGTTGTTCCCTTGCGGATCTTCTCCTTCTTCCCAGTTCCAGTCACAGTGAGGGCGCAT
CTCACCGGGTGGCTGATGACACTGAAGAAAACCTTCGTC

FIGURE 6

TGACGGAATCCCGGGCTGGGTATCCTGGTTTNGACAAGATAAACCCCCAGCAANAAATTGGG
GAGCAGGGCAAAACAGTNACGGGCAGCCCACATCAAGAAGTTCACCTTNGTTTGNATGGNTC
TGTCAACTCACGCTNTGTTTCGTGATGTTTTGGACACCCAAAGTGTTTGAGAAAATTTTGAT
AGACATNATCGGAGTGGANTTTGCCTTTGCAGAANTTTGNGNTGTTCCCTTTGCGGATTTTCT
CCTTTTTCCCAGTTCCAGTCACAGNGAGGGCGCATCTCACCGGGNGGNTGATGACANTGAAG
AAAACCTTTGTCCTTGCCCCCAGCTNTTTGGTGCGGATCATTGTCCTNATNGCCAGCCTTGT
GGTCCTACCCTACCTGGGGGTGCACGGTGCGACCCTGGGCGTGGGTTCCTCCTGGCGGGCA

FIGURE 7

TATTCCCAGTTCCGGTCACGGGGAGGGCGCATNTCACCGGGTGGCTGANGACACTGAAGAAA
ACCTTNGTCCTTGCCCCCAGNTTTGTGNTGCGGATNATCGTCCTCATCGCCAGCCTNGTGGT
CCTACCCTACCTGGGGGTGCACGGTGAGAC

Top of 23000

FIGURE 8

GCCCCGCGCCCGGCGCCGGGCGCCCGAAGCCGGGAGCCACCGCCATGGGGGCCTGCCTGGGA
GCCTGCTCCCTGCTCAGCTGCGCGTCCTGCCTCTGCGGCTCTGCCCCCTGCATCCTGTGCAG
CTGCTGCCCCGCCAGCCGCAACTCCACCGTGAGCCGCCTCATCTTCACGTTCTTCCTCTTCC
TGGGGGTGCTGGTGTCCATCATTATGCTGAGCCCGGGCGTGGAGAGTCAGCTCTACAAGCTG
CCCTGGGTGTGTGAGGAGGGGGCCGGGATCCCCACCGTCCTGCAGGGCCACATCGACTGTGG
CTCCCTGCTTGGCTACCGCGCTGTCTACCGCATGTGCTTCGCCACGGCGGCCTTCTTCTTCT
TCTTTTTTACCCCTGCTCATGCTCTGCGTGAGCAGCAGCCGGGACCCCCGGGCTGCCATCCAG
AATGGGTTTTGGTTCTTTAAGTTCCTGATCCTGGTGGGCCTCACCGTGGGTGCCTTCTACAT
CCCTGACGGCTCCTTCACCAACATCTGGTTCTACTTCGGCGTCGTGGGCTCCTTCCTCTTCA
TCCTCATCCAGCTGGTGTGCTCATCGACTTTGCGCACTCCTGGAACCAGCGGTGGCTGGGC
AAGGCCGAGGAGTGCGATTCCCGTGCCTGGTACGCAGGCCTCTTCTTCTTCACTCTCCTCTT
CTACTTGCTGTGATCGCGGCCGTGGCGCTGATGTTTATGTACTACACTGAGCCCAGCGGCT
GCCACGAGGGCAAGGTCTTCATCAGCCTCAACCTCACCTTCTGTGTCTGCGTGTCCATCGCT
GCTGTCTTGCCCAAGGTCCAGGACGCCCAGCCCAACTCGGGTCTGCTGCAGGCCTCGGTCAT
CACCTCTACACCATGTTTGTACCTGGTCAGCCCTATCCAGTATCCCTGAACAGAAATGCA
ACCCCCATTTGCCAACCAGCTGGGCAACGAGACAGTTGTGGCAGGCCCCGAGGGCTATGAG
ACCCAGTGGTGGGATGCCCCGAGCATTGTGGGCCTCATCATCTTCCTCCTGTGCACCCTCTT
CATCAGTCTGCGCTCCTCAGACCACCGGCAGGTGAACAGCCTGATGCAGACCGAGGAGTGCC
CACCTATGCTAGACGCCACACAGCAGCAGCAGCAGGTTGGCAGCCTGTGAGGGCCGGGCC
TTTGACAACGAGCAGGACGGCGTCACCTACAGCTACTCCTTCTTCCACTTCTGCCTGGTGCT
GGCCTCACTGCACGTCATGATGACGCTCACCAACTGGTACAAGCCCGGTGAGACCCGGAAGA
TGATCAGCACGTGGACCGCCGTGTGGGTGAAGATCTGTGCCAGCTGGGCAGGGCTGCTCCTC
TACCTGTGGACCCTGGTAGCCCCACTCCTCCTGCGCAACCGCGACTTCAGCTTGAGGGCAGCCT
CACAGCCTGCCATCTGGTGCCTCCTGCCACCTGGTGCCTCTCGGCTCGGTGACAGCCAACCT
GCCCCCTCCCCACACCAATCAGCCAGGCTGAGCCCCCACCCTGCCCCAGCTCCAGGACCTG
CCCCTGAGCCGGGCCTTCTAGTCGTAGTGCCTTCAGGGTCCGAGGAGCATCAGGCTCCTGCA
GAGCCCCATCCCCCGCCACACCCACACGGTGGAGCTGCCTCTTCCTTCCCCTCCTCCCTGT
TGCCCATACTCAGCATCTCGGATGAAAGGGCTCCCTTGTCTCCTCAGGCTCCACGGGAGCGGGG
CTGCTGGAGAGAGCGGGGAACCTCCACCCACAGTGGGGCATCCGGCACTGAAGCCCTGGTGTT
CCTGGTCACGTCCCCCAGGGGACCCTGCCCCCTTCTGGACTTCGTGCCTTACTGAGTCTCT
AAGACTTTTTCTAATAAAACAAGCCAGTGCGTGTAACAAAAA

FIGURE 9

MGACLGACSLSCASCLCGSAPCILCSCCPASRNSTVSRIFTFFLFLGVLVSIIMLSPGVE
SOLYKLPWVCEEAGIPTVLQGHIDCGSLLGYRAVYRMCFATAAFFFFFFFFTLLMLCVSSSRD
PRAAIQNGFWFFKFLILVGLTVGAFYIPDGSFTNIWFYFGVVGSLFILIQLVLLIDFAHSW
NQRWLKGAEECDSRAWYAGLFFFTLLFYLLSIAAVALMFMYYTEPSGCHEGKVFISLNLTF
VCVSIAAVLPKVQDAQPNSSGLLQASVITLYTMFVTWSALSSIQKCNPHLPTQLGNETVVA
GPEGYETQWWDAPSIVGLIIFLLCTLFISLRSSDHRQVNSLMQTEECPPMLDATQQQQQQVA
ACEGRAFDNEQDGVITYSYFFHFCLVLASLHVMMTLTNWYKPGETRKMISTWTAVVWKICAS
WAGLLLYLWTLVAPLLLRNRDFS

Signal sequence:

amino acids 1-20

Transmembrane domains:

amino acids 40-58, 101-116, 134-150, 162-178, 206-223, 240-257,
272-283, 324-340, 391-406, 428-444

FIGURE 10

GAGCGAGGCCGGGGACTGAAGGTGTGGGTGTGAGCCCTCTGGCAGAGGGTTAACCTGGGTC
AAATGCACGGATTCTCACCTCGTACAGTTACGCTCTCCCGCGGCACGTCCGCGAGGACTTGA
AGTCCTGAGCGCTCAAGTTTGTCCGTAGGTGAGAGAAGGCCATGGAGGTGCCGCCACCGGC
ACCGCGGAGCTTTCTCTGTAGAGCATTTGTGCCTATTTCCCCGAGTCTTTGCTGCCGAAGCTG
TGACTGCCGATTTCGGAAGTCCTTGAGGAGCGTCAGAAGCGGCTTCCCTACGTCCCAGAGCCC
TATTACCCGGAATCTGGATGGGACCGCCTCCGGGAGCTGTTTGGCAAAGATGAACAGCAGAG
AATTTCAAAGGACCTTGCTAATATCTGTAAGACGGCAGCTACAGCAGGCATCATTGGCTGGG
TGTATGGGGGAATACCAGCTTTTATTCATGCTAAACAACAATACATTGAGCAGAGCCAGGCA
GAAATTTATCATAACCGGTTTGATGCTGTGCAATCTGCACATCGTGCTGCCACACGAGGCTT
CATTCGTTATGGCTGGCGCTGGGGTTGGAGAACTGCAGTGTTTGTGACTATATTCAACACAG
TGAACACTAGTCTGAATGTATACCGAAATAAAGATGCCTTAAGCCATTTTGTAAATTGCAGGA
GCTGTCACGGGAAGTCTTTTTAGGATAAACGTAGGCCTGCGTGGCCTGGTGGCTGGTGGCAT
AATTGGAGCCTTGCTGGGCACTCCTGTAGGAGGCCTGCTGATGGCATTTCAGAAGTACGCTG
GTGAGACTGTTTCAGGAAAGAAAACAGAAGGATCGAAAGGCACTCCATGAGCTAAAACTGGAA
GAGTGGAAGGCAGACTACAAGTTACTGAGCACCTCCCTGAGAAAATTGAAAGTAGTTTACG
GGAAGATGAACCTGAGAATGATGCTAAGAAAATTGAAGCACTGCTAAACCTTCCTAGAAACC
CTTCAGTAATAGATAAACAAGACAAGGACTTGAAAGTGCTCTGAACTTGAACTCACTGGAGA
GCTGAAGGGAGCTGCCATGTCCGATGAATGCCAACAGACAGGCCACTCTTTGGTCAGCCTGC
TGACAAATTTAAGTGCTGGTACCTGTGGTGGCAGTGGCTTGCTCTTGCTCTTTTTCTTTCTT
TTTAACTAAGAATGGGGCTGTTGTACTCTCACTTTACTTATCCTTAAATTTAAATACATACT
TATGTTTGTATTAATCTATCAATATATGCATACATGGATATATCCACCCACCTAGATTTTAA
GCAGTAAATAAAACATTTTCGCAAAAGATTAAAGTTGAATTTTACAGTTT

FIGURE 11

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA23318

><subunit 1 of 1, 285 aa, 1 stop

><MW: 32190, pI: 9.03, NX(S/T): 2

MEVPPPAPRSFLCRALCLFPRVFAAEAVTADSEVLEERQKRLPYVPEPYYPESGWDRLRELF
GKDEQQRISKDLANICKTAATAGIIGWVYGGIPAFIHAKQQYIEQSQAIEYHNRFDVQSAH
RAATRGFIRYGWRWGWRTAVFVTIFNTVNTSLNVYRNKDALSHFVIAGAVTGSLFRINVGLR
GLVAGGIIGALLGTPVGGLLMAFQKYAGETVQERKQKDRKALHELKLEEWKGRLQVTEHLPE
KIESSLREDEPENDAKKIEALLNLPRNPSVIDKQDKD

Important Features:

Signal Peptide:

amino acids 1-24

Transmembrane domains:

amino acids 76-96 and 171-195

N-glycosylation site:

amino acids 153-156

FIGURE 12

CGGAAGTCCCTTGAGGAGCGTCAGAAGCGGCTTCCCTACGTCCCAGAGCCCTATTACCCGGA
ATCTGGATGGGACCGCTCCGGGAGCTGTTTGGCAAAGATGAACAGCAGAGAATTTCAAAGGA
CCTTGCTAATATCTGTAAGACGGCAGCTACAGCAGGCATCATTGGCTGGGTGTATGGGGGAA
TACCAGCTTTTATTCATGCTAAACAACAATACATTGAGCAGAGCCAGGCAGAAATTTATCAT
AACCGGTTTGATGCTGTGCAATCTGCACATCGTGCTGCCACACGAGGCTTCATTCGTTTCATG
GCTGGCGCCGAACC

FIGURE 13

TCAAGTTTGTCCGTAGGTCGAGAGAAGGCCATGGAGGTGCCGCCACCGGCACCGCGGAGCTT
TTTTCTGTAGAGCATTGTGCCTATTTCCCCGAGTTTTTGCTGCCGAAGCTGTGACTGCCGAT
TCGGAAGTCCTTGAGGAGCGTCAGAAGCGGCTTCCCTACGTCCCAGAGCCCTATTACCCGGA
ATTTGGATGGGACCGCCTCCGGGAGCTGTTTGGCAAAGATGAACAGCAGAGAATTTCAAAGG
ACCTTGCTGATATNTGTAAGACGGCAGCTACAGCAGGCATCATTGGCTGGGTGTATGGGGGA
ATACCAGCTTTTATTTCATGNTAAACAACAATACATTGAGCAGAGCCAGGCAGAAATTTATNA
TAACC

[illegible]

GAGCCGCGCCGCGCGCGCGCACTGCAGCCCCAGGCCCGCCGCCACCCACGTCT
GCGTTGCTGCCCCGCCTGGGCCAGGCCCCAAAGGCAAGGACAAAGCAGCTGTCAGGGAACCT
CCGCCGGAGTCAAAATTTACGTGCAGCTGCCGGCAACCACAGGTTCCAAGATGGTTTTCGGGG
GCTTCGCGTGTTCCAAGAACTGCCTGTGCGCCCTCAACCTGCTTTACACCTTGGTTAGTCTG
CTGCTAATTGGAATTGCTGCGTGGGGCATTGGCTTCGGGCTGATTTCCAGTCTCCGAGTGGT
CGGCGTGGTCATTGCAGTGGGCATCTTCTTGTTCTGATTGCTTTAGTGGGTCTGATTGGAG
CTGTAAACATCATCAGGTGTTGCTATTTTTTTTATATGATTATTCTGTTACTTGTATTTATT
GTTTCAGTTTTCTGTATCTTGCGCTTGTTTAGCCCTGAACCAGGAGCAACAGGGTCAGCTTCT
GGAGGTTGGTTGGAACAATACGGCAAGTGCTCGAAATGACATCCAGAGAAATCTAAACTGCT
GTGGGTTCCGAAGTGTTAACCCAAATGACACCTGTCTGGCTAGCTGTGTTAAAAGTGACCAC
TCGTGCTCGCCATGTGCTCCAATCATAGGAGAATATGCTGGAGAGGTTTTGAGATTTGTTGG
TGGCATTGGCCTGTTCTTCAGTTTTACAGAGATCCTGGGTGTTTGGCTGACCTACAGATACA
GGAACCAGAAAGACCCCCGCGCGAATCCTAGTGCAATCCTTTTGATGAGAAAAAAGGAAGAT
TTCTTTTCGTATTATGATCTTGTTCACTTTCTGTAATTTTCTGTTAAGCTCCATTTGCCAGT
TTAAGGAAGGAAACACTATCTGGAAAAGTACCTTATTGATAGTGGAATTATATATTTTTTACT
CTATGTTTTCTCTACATGTTTTTTTTCTTTCCGTTGCTGAAAAATATTTGAAACTTGTGGTCTC
TGAAGCTCGGTGGCACCTGGAATTTACTGTATTCATTGTCGGGCACTGTCCACTGTGGCCTT
TCTTAGCATTTTTTACCTGCAGAAAACTTTGTATGGTACCACTGTGTTGGTTATATGGTGAA
TCTGAACGTACATCTCACTGGTATAATTATATGTAGCACTGTGCTGTGTAGATAGTTCCTAC
TGGAAAAAGAGTGGAATTTATTAAAAATCAGAAAGTATGAGATCCTGTTATGTTAAGGGAAA
TCCAAATTCCCAATTTTTTTTTGGTCTTTTTTAGGAAAGATTGTTGTGGTAAAAAGTGTTAGTA
TAAAAATGATAATTTACTTGTAGTCTTTTATGATTACACCAATGTATTCTAGAAATAGTTAT
GTCTTAGGAAATTGTGGTTTAATTTTTTGACTTTTACAGGTAAGTGCAAAGGAGAAGTGGTTT
CATGAAATGTTCTAATGTATAATAACATTTACCTTCAGCCTCCATCAGAATGGAACGAGTTT
TGAGTAATCAGGAAGTATATCTATATGATCTTGATATTGTTTTATAATAATTTGAAGTCTAA
AAGACTGCATTTTTTAAACAAGTTAGTATTAATGCGTTGGCCACGTAGCAAAAAGATATTTG
ATTATCTTAAAAATTGTTAAATACCGTTTTTCATGAAATTTCTCAGTATTGTAACAGCAACTT
GTCAAACCTAAGCATATTTGAATATGATCTCCATAATTTGAAATTGAAATCGTATTGTGTG
GCTCTGTATATTCTGTTAAAAAATTAAAGGACAGAAACCTTTCTTTGTGTATGCATGTTTGA
ATTAAAAGAAAGTAATGGAAG

FIGURE 15

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA39979

><subunit 1 of 1, 204 aa, 1 stop

><MW: 22147, pI: 8.37, NX(S/T): 3

MVCGGFACSKNCLCALNLLYTLVSLLLIGIAAWGIGFGLISSLRVVGVIKAVGIFLFLIALV
GLIGAVKHHQVLLFFYMIILLLVFIVQFSVSCACLALNQEQGQLLEVGNNTASARNDIQR
NLNCCGFRSVNPNDTCLASCVKSDHSCSPCAPIIGEYAGEVLRVFGGIGLFFSFTEILGVWL
TYRYRNQKDPRANPSAFL

Signal Peptide:

amino acids 1-34

Transmembrane domains:

amino acids 47-63, 72-95 and 162-182

FIGURE 16

TGATTGGAGCTGTAAAAAANTCTTCAGGTGTTGTNATTTTTTTTATATGATTATTCTGTAANT
TGTATTTTATTGTTTTCAGTTTTNTGTATCTTGCGCTTGTTTAGCCNTGAACCAGGAGCAACAGG
GTCAGNTTNTGGAGGTTGGTTGGAACAATACGGCAAGTGCTCGAAATGACATCCAGAGAAAT
NTAAACTGCTGTGGGTTCCGAAGTGTTAACCCTAAATGACACCTGTNTGGCTAGCTGTGTTAA
AAGTGACCACTNGTGCTCGCCATGTGCTCCAATCATAGGAGAATATGCTGGAGAGGTTTTGA
GATTTGTTGGTGGCATTGGCCTGTTNTTCAGTTTTACAGAGATCCTGGGTGTTTGGCTGACC
TACAGATACAGGAACCAG

FIGURE 17

AATCCCAAATTCCCAATTTTTTTGGNCTTTTTAGGGAAAGATGTGTTGTGGTAAAAAGTGT
TAGTATAAAAATGATAATTTACTTGTAGTCTTTTATGATTACACCAATGTATTCTAGAATAG
TTATGTCTTAGGAAATTGTGGTTTAATTTTTGACTTTTACAGGTAAGTGCAAAGGAGAAGTG
GTTTCATGAAATGTTCTAATGTATAATAACATTTACCTTCAGCCTCCCATCAGAATGGAACG
AGTTTTGAGTAATCCAGGAAGTATATCTATATGATCTTGATATTGTTTTATATAATTTGAAG
TCTAAAAGACTGCATTTTTTAAACAAGTTAGTATTAATGCGTTGGCCACGTAGCAAAAAGAT
ATTTGATTATCTTAAAAATTGTTAAATACCGTTTTTCATGAAAGTTCTCAGTATTGTAACAGC
AACTTGTCAAACCTAAGCATATTTGAATATGATCTCCATAATTTGAAATTGAAATCGTATT
GTGTGGAGGAAATGGCAATCTTATGTGTGCTGAAGGACACAGTAAGAGCACCAAGTTGTGCC
CCTTTGC

FIGURE 18

ATGATTATTCTGTTACTTGTATTTATTGTTTCAGTTTTATGGTATCTTGCGCTTGTTTAGCCC
CTGAAACCAGGAGCAACAGGGNNCAGCTTCCTGGAGGTTGGTTGGCAACAATCACGGCCAAG
TGACTCCGCAAATGACATCCCAGAGAAATCCTAAACTGCTGTGGGTTCCGAAGTGTTAACCC
AAATGACACCTGTCTGGCTNGCTGTGTTAAAAGTGACCACTCGTGCTCGCCATGTGCTCCAA
TCATAGGAGAATATGC

FIGURE 19

CAGTCACCATGAAGCTGGGCTGTGTCCTCATGGCCTGGGCCCTCTACCTTTCCCTTGGTGTG
CTCTGGGTGGCCCAGATGCTACTGGCTGCCAGTTTTTGAGACGCTGCAGTGTGAGGGACCTGT
CTGCACTGAGGAGAGCAGCTGCCACACGGAGGATGACTTGACTGATGCAAGGGAAGCTGGCT
TCCAGGTCAAGGCCTACACTTTCAGTGAACCCTTCCACCTGATTGTGTCCTATGACTGGCTG
ATCCTCCAAGGTCCAGCCAAGCCAGTTTTTGAAGGGGACCTGCTGGTTCTGCGCTGCCAGGC
CTGGCAAGACTGGCCACTGACTCAGGTGACCTTCTACCGAGATGGCTCAGCTCTGGGTCCCC
CCGGGCCTAACAGGGAATTCTCCATCACCGTGGTACAAAAGGCAGACAGCGGGCACTACCAC
TGCAGTGGCATCTTCCAGAGCCCTGGTCTGGGATCCCAGAAACAGCATCTGTTGTGGCTAT
CACAGTCCAAGAACTGTTTCCAGCGCCAATTCTCAGAGCTGTACCCTCAGCTGAACCCCAAG
CAGGAAGCCCCATGACCCTGAGTTGTGAGACAAAGTTGCCCTGCAGAGGTCAGCTGCCCCG
CTCCTCTTCTCCTTCTACAAGGATGGAAGGATAGTGCAAAGCAGGGGGCTCTCCTCAGAATT
CCAGATCCCCACAGCTTCCAGAAGATCACTCCGGGTCTACTGGTGTGAGGCAGCCACTGAGG
ACAACCAAGTTTGGAAACAGAGCCCCCAGCTAGAGATCAGAGTGCAGGGTGCTTCCAGCTCT
GCTGCACCTCCACATTGAATCCAGCTCCTCAGAAATCAGCTGCTCCAGGAACTGCTCCTGA
GGAGGCCCCCTGGGCCTCTGCCTCCGCCGCAACCCCATCTTCTGAGGATCCAGGCTTTTCTT
CTCCTCTGGGGATGCCAGATCCTCATCTGTATCACCAGATGGGCCTTCTTCTCAAACACATG
CAGGATGTGAGAGTCCTCCTCGGTCACTGTCTCATGGAGTTGAGGGAATTATCTGGCCACCA
GAAGCCTGGGACCACAAAGGCTACTGCTGAATAGAAGTAAACAGTTTCATCCATGATCTCACT
TAACCACCCCAATAAATCTGATTCTTTATTTTCTCTTCTGTCTGCACATATGCATAAGTA
CTTTTACAAGTTGTCCCACTGTTTTGTGTTAGAATAATGTAGTTAGGTGAGTGTAAATAAATTT
ATATAAAGTGAGAATTAGAGTTTAGCTATAAATTGTGTATTCTCTCTTAACACAACAGAATTC
TGCTGTCTAGATCAGGAATTTCTATCTGTTATATCGACCAGAATGTTGTGATTTAAAGAGAA
CTAATGGAAGTGGAATTGAATACAGCAGTCTCAACTGGGGGCAATTTTGCCCCCAGAGGACA
TTGGGCAATGTTTGGAGACATTTTGGTCATTATACTTGGGGGGTTGGGGGATGGTGGGATGT
GTGTCTACTGGCATCCAGTAAATAGAAGCCAGGGGTGCCGCTAAACATCCTATAATGCACAG
GGCAGTACCCCAACGAAAAATAATCTGGCCCCAAATGTGAGTTGTAAGTGTGAGTTTGAAGAA
CCCCAGCCTAATGAAACCCTAGGTGTTGGGCTCTGGAATGGGACTTTGTCCCTTCTAATTAT
TATCTCTTTCCAGCCTCATTGAGCTATTCTTACTGACATACCAGTCTTTAGCTGGTGTCTATG
GTCTGTTCTTTAGTTCTAGTTTGTATCCCCTCAAAGGCCATTATGTTGAAATCCTAATCCCC
AAGGTGATGGCATTAAGAAGTGGGCCTTTGGGAAGTGATTAGATCAGGAGTGCAGAGCCCTC
ATGATTAGGATTAGTGCCCTTATTTAAAAAGGCCCCAGAGAGCTAACTCACCTTCCACCAT
ATGAGGACGTGGCAAGAAGATGACATGTATGAGAACCAAAAAACAGCTGTGCCAAACACCG
ACTCTGTCGTTGCCTTGATCTTGAACCTCCAGCCTCCAGAACTATGAGAAATAAAATTCTGG
TTGTTTGTAGCCTAA

FIGURE 20

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA40594

><subunit 1 of 1, 359 aa, 1 stop

><MW: 38899, pI: 5.21, NX(S/T): 0

MKLGCVLMAWALYLSLGV LWVAQMLLAASFETLQCEGPVCTEESCHTEDDLTDAREAGFQV
KAYTFSEPFHLIVSYDWLILQGPAKPVFEGDLLVLRQAWQDWPLTQVTFYRDGSALGPPGP
NREFSITVVQKADSGHYHCSGIFQSPGPGIPETASVVAITVQELFPAPILRAVPSAEPQAGS
PMTLSCQTKLPLQRSAARLLFSFYKDGRIVQSRGLSSEFQIPTASEDHSGSYWCEAATEDNQ
VWKQSPQLEIRVQGASSAAPTILNPAPQKSAAPGTAPPEEAPGPLPPPPTPSSSEDPGFSSPL
GMPDPHLYHQMGLLLKHMQDVRVLLGHLLMELRELSGHQKPGTTKATAE

Signal sequence:

amino acids 1-17

Leucine zipper pattern sequence:

amino acids 12-33

Protein kinase C phosphorylation site:

amino acids 353-355

FIGURE 21

CCCACGCGTCCGCCCACGCGTCCGCCCACGGGTCCGCCCACGCGTCCGGGCCACCAGAAGTT
TGAGCCTCTTTGGTAGCAGGAGGCTGGAAGAAAGGACAGAAGTAGCTCTGGCTGTGATGGGG
ATCTTACTGGGCCTGCTACTCCTGGGGCACCTAACAGTGGACACTTATGGCCGTCCCATCCT
GGAAGTGCCAGAGAGTGTAAACAGGACCTTGGAAGGGGATGTGAATCTTCCCTGCACCTATG
ACCCCTGCAAGGCTACACCCAAGTCTTGGTGAAGTGGCTGGTACAACGTGGCTCAGACCCT
GTCACCATCTTTCTACGTGACTCTTCTGGAGACCATATCCAGCAGGCAAAGTACCAGGGCCG
CCTGCATGTGAGCCACAAGGTTCCAGGAGATGTATCCCTCCAATTGAGCACCTTGGAGATGG
ATGACCGGAGCCACTACACGTGTGAAGTCACCTGGCAGACTCCTGATGGCAACCAAGTCGTG
AGAGATAAGATTACTGAGCTCCGTGTCCAGAACTCTCTGTCTCCAAGCCCACAGTGACAAC
TGGCAGCGTTATGGCTTCACGGTGCCCCAGGGAATGAGGATTAGCCTTCAATGCCAGGCTC
GGGGTTCTCCTCCCATCAGTTATATTTGGTATAAGCAACAGACTAATAACCAGGAACCCATC
AAAGTAGCAACCCTAAGTACCTTACTCTTCAAGCCTGCGGTGATAGCCGACTCAGGCTCCTA
TTTCTGCACTGCCAAGGGCCAGGTTGGCTCTGAGCAGCACAGCGACATTGTGAAGTTTGTGG
TCAAAGACTCCTCAAAGCTACTCAAGACCAAGACTGAGGCACCTACAACCATGACATACCCC
TTGAAAGCAACATCTACAGTGAAGCAGTCCTGGGACTGGACCACTGACATGGATGGCTACCT
TGGAGAGACCAGTGCTGGGCCAGGAAAGAGCCTGCCTGTCTTTGCCATCATCCTCATCATCT
CCTTGTGCTGTATGGTGGTTTTTACCATGGCCTATATCATGCTCTGTGCGAAGACATCCCAA
CAAGAGCATGTCTACGAAGCAGCCAGGTAAGAAAGTCTCTCCTCTTCCATTTTTTGACCCCGT
CCCTGCCCTCAATTTTTGATTACTGGCAGGAAATGTGGAGGAAGGGGGGTGTGGCACAGACCC
AATCCTAAGGCCGGAGGCCTTCAGGGTCAGGACATAGCTGCCTTCCCTCTCTCAGGCACCTT
CTGAGGTTGTTTTGGCCCTCTGAACACAAAGGATAATTTAGATCCATCTGCCTTCTGCTTCC
AGAATCCCTGGGTGGTAGGATCCTGATAATTAATTGGCAAGAATTGAGGCAGAAGGGTGGGA
AACCAGGACCACAGCCCCAAGTCCCTTCTTATGGGTGGTGGGCTCTTGGGCCATAGGGCACA
TGCCAGAGAGGCCAACGACTCTGGAGAAACCATGAGGGTGGCCATCTTCGCAAGTGGCTGCT
CCAGTGATGAGCCAACTTCCCAGAATCTGGGCAACAACACTACTCTGATGAGCCCTGCATAGGA
CAGGAGTACCAGATCATCGCCAGATCAATGGCAACTACGCCCGCCTGCTGGACACAGTTCC
TCTGGATTATGAGTTTCTGGCCACTGAGGGCAAAAGTGTCTGTTAAAAATGCCCCATTAGGC
CAGGATCTGCTGACATAATTGCCTAGTCAGTCCTTGCCTTCTGCTATGGCCTTCTTCCCTGCT
ACCTCTCTTCCCTGGATAGCCCAAAGTGTCCGCTACCAACACTGGAGCCGCTGGGAGTCACT
GGCTTTGCCCTGGAATTTGCCAGATGCATCTCAAGTAAGCCAGCTGCTGGATTTGGCTCTGG
GCCCTTCTAGTATCTCTGCCGGGGGCTTCTGGTACTCCTCTCTAAATACCAGAGGGAAGATG
CCCATAGCACTAGGACTTGGTCATCATGCCTACAGACACTATTCAACTTTGGCATCTTGCCA
CCAGAAGACCCGAGGGAGGCTCAGCTCTGCCAGCTCAGAGGACCAGCTATATCCAGGATCAT
TTCTCTTTCTTCAGGGCCAGACAGCTTTTAATTGAAATTGTTATTTACAGGCCAGGGTTCA
GTTCTGCTCCTCCACTATAAGTCTAATGTTCTGACTCTCTCCTGGTGCTCAATAAATATCTA
ATCATAACAGC

FIGURE 22

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA45416

><subunit 1 of 1, 321 aa, 1 stop

><MW: 35544, pI: 8.51, NX(S/T): 0

MGILLGLLLLGHLLTVDTYGRPILEVPESVTGPWKGDVNLPCYDPLQGYTQVLVKWLVQRGS
DPVTIFLRDSSGDHIQQAKYQGRLLHVSHKVPGDVSLQLSTLEMDDRSHYTCEVTWQTPDGNQ
VVRDKITELRVQKLSVSKPTVTTGSGYGFTVPQGMRLSLQCQARGSPPIISYIWYKQQTNNQE
PIKVATLSTLLFKPAVIADSGSYFCTAKGQVGSEQHSDIVKFVVKDSSKLLKTKTEAPTTMT
YPLKATSTVKQSWDWTDDMDGYLGETSAGPGKSLPVFAIILIISLCCMVVFTMAYIMLCRKT
SQQEHVYEAAAR

Signal Sequence:

amino acids 1-19

Glycosaminoglycan attachment site:

amino acids 149-152

Transmembrane domain:

amino acids 282-300

FIGURE 23

GCGCCGGGAGCCCATCTGCCCCCAGGGGACGCGGGGCGCGGGGCCGGCTCCCGCCCGGCACAT
GGCTGCAGCCACCTCGCGCGCACCCCGAGGCGCCGCGCCAGCTCGCCCGAGGTCCGTCCGA
GGCGCCCGGCGCCCGCCCGGAGCCAAGCAGCAACTGAGCGGGGAAGCGCCCGCGTCCGGGGATC
GGG**ATG**TCCCTCCTCCTTCTCCTCTTGCTAGTTTCCTACTATGTTGGAACCTTGGGGACTCA
CACTGAGATCAAGAGAGTGGCAGAGGAAAAGGTCACTTTGCCCTGCCACCATCAACTGGGGC
TTCCAGAAAAAGACACTCTGGATATTGAATGGCTGCTCACCGATAATGAAGGGAACCAAAAA
GTGGTGATCACTTACTCCAGTCGTCTGTCTACAATAACTTGACTGAGGAACAGAAGGGCCG
AGTGGCCTTTGCTTCCAATTTCTGGCAGGAGATGCCTCCTTGAGATTGAACCTCTGAAGC
CCAGTGATGAGGGCCGGTACACCTGTAAGGTTAAGAATTCAGGGCGCTACGTGTGGAGCCAT
GTCATCTTAAAAGTCTTAGTGAGACCATCCAAGCCCAAGTGTGAGTTGGAAGGAGAGCTGAC
AGAAGGAAGTGACCTGACTTTGTCAGTGTGAGTCATCCTCTGGCACAGAGCCCATTTGTGTATT
ACTGGCAGCGAATCCGAGAGAAAGAGGGAGAGGATGAACGTCTGCCTCCCAAATCTAGGATT
GACTACAACCACCCCTGGACGAGTTCTGCTGCAGAATCTTACCATGTCTACTCTGGACTGTA
CCAGTGACACAGCAGGCAACGAAGCTGGGAAGGAAAGCTGTGTGGTGCGAGTAACTGTACAGT
ATGTACAAAGCATCGGCATGGTTGCAGGAGCAGTGACAGGCATAGTGGCTGGAGCCCTGCTG
ATTTTCCTCTTGGTGTGGCTGCTAATCCGAAGGAAAGACAAAGAAAGATATGAGGAAGAAGA
GAGACCTAATGAAATTCGAGAAGATGCTGAAGCTCCAAAAGCCCGTCTTGTGAAACCCAGCT
CCTCTTCCTCAGGCTCTCGGAGCTCACGCTCTGGTTCTTCCTCCACTCGCTCCACAGCAAAT
AGTGCCTCACGCAGCCAGCGGACACTGTCAACTGACGCAGCACCCAGCCAGGGCTGGCCAC
CCAGGCATACAGCCTAGTGGGGCCAGAGGTGAGAGGTTCTGAACCAAAGAAAGTCCACCATG
CTAATCTGACCAAAGCAGAAACCACACCCAGCATGATCCCCAGCCAGAGCAGAGCCTTCCAA
ACGGTCT**TGA**ATTACAATGGACTTGACTCCCACGCTTTCTTAGGAGTCAGGGTCTTTGGACTC
TTCTCGTCATTGGAGCTCAAGTCACCAGCCACACAACCAGATGAGAGGTCATCTAAGTAGCA
GTGAGCATTGCACGGAACAGATTCAGATGAGCATTTTTCCTTATACAAATACCAAAACAAGCAA
AGGATGTAAGCTGATTCTGTAAAAAGGCATCTTATTGTGCCTTTAGACCAGAGTAAGGG
AAAGCAGGAGTCCAAATCTATTTGTTGACCAGGACCTGTGGTGAGAAGGTTGGGGAAAGGTG
AGGTGAATATACCTAAAACTTTTAATGTGGGATATTTTGTATCAGTGCTTTGATTACAAATT
TTCAAGAGGAAATGGGATGCTGTTTGTAATTTTCTATGCATTTCTGCAAACCTTATTGGATT
ATTAGTTATTACAGACAGTCAAGCAGAAACCCACAGCCTTATTACACCTGTCTACACCATGTAC
TGAGCTAACCCTTCTAAGAACTCCAAAAAAGGAAACATGTGTCTTCTATTCTGACTTAAC
TTCATTTGTCATAAGGTTTGGATATTAATTTCAAGGGGAGTTGAAATAGTGGGAGATGGAGA
AGAGTGAATGAGTTTCTCCCACTCTATACTAATCTCACTATTTGTATTGAGCCCAAAATAAC
TATGAAAGGAGACAAAAATTTGTGACAAAGGATTGTGAAGAGCTTTCCATCTTCATGATGTT
ATGAGGATTGTTGACAAACATTAGAAATATATAATGGAGCAATTGTGGATTTCCCCTCAAAT
CAGATGCCTCTAAGGACTTTCCTGCTAGATATTTCTGGAAGGAGAAAAATACAACATGTCATT
TATCAACGTCCTTAGAAAGAATCTTCTAGAGAAAAAGGGATCTAGGAATGCTGAAAGATTA
CCCAACATACCATTATAGTCTCTTCTTTCTGAGAAAATGTGAAACCAGAATTGCAAGACTGG
GTGGACTAGAAAGGGAGATTAGATCAGTTTTCTCTTAATATGTCAAGGAAGGTAGCCGGGCA
TGGTGCCAGGCACCTGTAGGAAAATCCAGCAGGTGGAGGTTGCAGTGAGCCGAGATTATGCC
ATTGCACTCCAGCCTGGGTGACAGAGCGGGACTCCGTCTC

FIGURE 24

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA45419

><subunit 1 of 1, 373 aa, 1 stop

><MW: 41281, pI: 8.33, NX(S/T): 3

MSLLLLLLLLVSYYVGTLGTHTEIKRVAEEKVTLPCHHQLGLPEKDTLDIEWLLTDNEGNQKV
VITYSSRHVYNNLTEEQKGRVAFASNFLAGDASLQIEPLKPSDEGRYTCKVKNSGRYVWSHV
ILKVLVRPSKPKCELEGELTEGSDLTLCESSSGTEPIVYYWQRIREKEGEDERLPPKSRID
YNHPGRVLLQNLTMSYSGLYQCTAGNEAGKESCVVRVTVQYVQSIGMVAGAVTGIVAGALLI
FLLVWLLIRRKDKERYEEEEERPNEIREDAEAPKARLVKPSSSSSSGSRSSSRSGSSSTRSTANS
ASRSQRTLSTDAAPQPGLATQAYSLVGPEVRGSEPKKVHHANLTKAETTPSMIPSQSRAFQTV

Signal sequence:

amino acids 1-16

Transmembrane domain:

amino acids 232-251

	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030	2031	2032	2033	2034	2035	2036	2037	2038	2039	2040	2041	2042	2043	2044	2045	2046	2047	2048	2049	2050	2051	2052	2053	2054	2055	2056	2057	2058	2059	2060	2061	2062	2063	2064	2065	2066	2067	2068	2069	2070	2071	2072	2073	2074	2075	2076	2077	2078	2079	2080	2081	2082	2083	2084	2085	2086	2087	2088	2089	2090	2091	2092	2093	2094	2095	2096	2097	2098	2099	2100	2101	2102	2103	2104	2105	2106	2107	2108	2109	2110	2111	2112	2113	2114	2115	2116	2117	2118	2119	2120	2121	2122	2123	2124	2125	2126	2127	2128	2129	2130	2131	2132	2133	2134	2135	2136	2137	2138	2139	2140	2141	2142	2143	2144	2145	2146	2147	2148	2149	2150	2151	2152	2153	2154	2155	2156	2157	2158	2159	2160	2161	2162	2163	2164	2165	2166	2167	2168	2169	2170	2171	2172	2173	2174	2175	2176	2177	2178	2179	2180	2181	2182	2183	2184	2185	2186	2187	2188	2189	2190	2191	2192	2193	2194	2195	2196	2197	2198	2199	2200	2201	2202	2203	2204	2205	2206	2207	2208	2209	2210	2211	2212	2213	2214	2215	2216	2217	2218	2219	2220	2221	2222	2223	2224	2225	2226	2227	2228	2229	2230	2231	2232	2233	2234	2235	2236	2237	2238	2239	2240	2241	2242	2243	2244	2245	2246	2247	2248	2249	2250	2251	2252	2253	2254	2255	2256	2257	2258	2259	2260	2261	2262	2263	2264	2265	2266	2267	2268	2269	2270	2271	2272	2273	2274	2275	2276	2277	2278	2279	2280	2281	2282	2283	2284	2285	2286	2287	2288	2289	2290	2291	2292	2293	2294	2295	2296	2297	2298	2299	2300	2301	2302	2303	2304	2305	2306	2307	2308	2309	2310	2311	2312	2313	2314	2315	2316	2317	2318	2319	2320	2321	2322	2323	2324	2325	2326	2327	2328	2329	2330	2331	2332	2333	2334	2335	2336	2337	2338	2339	2340	2341	2342	2343	2344	2345	2346	2347	2348	2349	2350	2351	2352	2353	2354	2355	2356	2357	2358	2359	2360	2361	2362	2363	2364	2365	2366	2367	2368	2369	2370	2371	2372	2373	2374	2375	2376	2377	2378	2379	2380	2381	2382	2383	2384	2385	2386	2387	2388	2389	2390	2391	2392	2393	2394	2395	2396	2397	2398	2399	2400	2401	2402	2403	2404	2405	2406
--	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------

AAAAAAAAAA

FIGURE 26

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA52594

><subunit 1 of 1, 655 aa, 1 stop

><MW: 71845, pI: 8.22, NX(S/T): 8

MGTSPSSSTALASCSRIARRATATMIAGSLLLLGFLSTTTAQPEQKASNLIGTYRHVD RATG
QVLTCDKCPAGTYVSEHCTNTSLRVCSSCPVGTFTRHENGIEKCHDCSQPCWP MIEKLPCA
ALTDRECTCPPGMFQSNATCAPHTVCPVGWGVRRKKGTETEDVRCKQCARGTFSDVPSSVMKC
KAYTDCLSQNLVVIKPGTKETDNVCGTLPSFSSSTSPSPGTAIFPRPEHMETHEVPSSTYVP
KGMNSTESNSSASVRPKVLSSIQEGTVPDNTSSARGKEDVNKTLPNLQVVNHQQGPHHRHIL
KLLPSMEATGGEKSSTPIKGPKRGHPRQNLHKHFDINEHLPWMIVLFLLLVLVVIVVCSIRK
SSRTLKKGPRQDPSAIVEKAGLKKSMPTQNREKWIYYCNGHGIDILKLVA AQVGSQWKDIY
QFLCNASEREVA AFSNGYTADHERAYAALQHW TIRGPEASLAQLISALRQHRNDVVEKIRG
LMEDTTQLETDKLALPMSPSPLSPSPIPSPNAKLENSALLTVEPSPQDKNKGFFVDESEPLL
RCDSTSSGSSALS RNGSFITKEKKD TVLRQVRLDPCDLQPIFDDMLHFLNPEELRVIEEIPQ
AEDKLDRLFEIIGVKSQEASQTLLDSVYSHLPDLL

Signal sequence:

amino acids 1-41

Transmembrane domain:

amino acids 350-370

FIGURE 27

ATGGGAAGCCAGTAACACTGTGGCCTACTATCTCTTCCGTGGTGCCATCTACATTTTTTGGGA
CTCGGGAATTATGAGGTAGAGGTGGAGGCGGAGCCGGATGTCAGAGGTCCTGAAATAGTCAC
CATGGGGGAAAATGATCCGCCTGCTGTTGAAGCCCCCTTCTCATTCCGATCGCTTTTTTGGCC
TTGATGATTTGAAAATAAGTCCTGTTGCACCAGATGCAGATGCTGTTGCTGCACAGATCCTG
TCACTGCTGCCATTGAAGTTTTTTTCCAATCATCGTCATTGGGATCATTGCATTGATATTAGC
ACTGGCCATTGGTCTGGGCATCCACTTCGACTGCTCAGGGAAGTACAGATGTCGCTCATCCT
TTAAGTGTATCGAGCTGATAGCTCGATGTGACGGAGTCTCGGATTGCAAAGACGGGGAGGAC
GAGTACCGCTGTGTCCGGGTGGGTGGTCAGAAATGCCGTGCTCCAGGTGTTACAGCTGCTTC
GTGGAAGACCATGTGCTCCGATGACTGGAAGGGTCACTACGCAAATGTTGCCTGTGCCAAC
TGGGTTTCCCAAGCTATGTGAGTTTCAGATAAACCTCAGAGTGAGCTCGCTGGAGGGGCAGTTC
CGGGAGGAGTTTGTGTCCATCGATCACCTCTTGCCAGATGACAAGGTGACTGCATTACACCA
CTCAGTATATGTGAGGGAGGGATGTGCCTCTGGCCACGTGGTTACCTTGACGTGCACAGCCT
GTGGTTCATAGAAGGGGCTACAGCTCACGCATCGTGGGTGGAAACATGTCCTTGCTCTCGCAG
TGGCCCTGGCAGGCCAGCCTTCAGTTCCAGGGCTACCACCTGTGCGGGGGCTCTGTATCAC
GCCCCTGTGGATCATCACTGCTGCACACTGTGTTTATGACTTGTACCTCCCCAAGTCATGGA
CCATCCAGGTGGGTCTAGTTTTCCCTGTTGGACAATCCAGCCCCATCCCACCTTGGTGGAGAAG
ATTGTCTACCACAGCAAGTACAAGCCAAAGAGGCTGGGCAATGACATCGCCCTTATGAAGCT
GGCCGGGCCACTCACGTTCAATGAAATGATCCAGCCTGTGTGCCTGCCCAACTCTGAAGAGA
ACTTCCCCGATGGAAAAGTGTGCTGGACGTCAGGATGGGGGGCCACAGAGGATGGAGGTGAC
GCCTCCCCGTCTCTGAACCACGCGGCCGTCCCTTTGATTTCCAACAAGATCTGCAACCACAG
GGACGTGTACGGTGGCATCATCTCCCCCTCCATGCTCTGCGCGGGCTACCTGACGGGTGGCG
TGGACAGCTGCCAGGGGGACAGCGGGGGGCCCTGGTGTGTCAAGAGAGGAGGCTGTGGAAG
TTAGTGGGAGCGACCAGCTTTGGCATCGGCTGCGCAGAGGTGAACAAGCCTGGGGTGTACAC
CCGTGTACCTCCTTCCCTGGACTGGATCCACGAGCAGATGGAGAGAGACCTAAAAACCTGAA
GAGGAAGGGGACAAGTAGCCACCTGAGTTCCCTGAGGTGATGAAGACAGCCCGATCCTCCCCCT
GGACTCCCGTGTAGGAACCTGCACACGAGCAGACACCCTTGGAGCTCTGAGTTCCGGCACCA
GTAGCAGGCCCCGAAAGAGGCACCCTTCCATCTGATTCCAGCACAACTTCAAGCTGCTTTTT
GTTTTTTGTTTTTTTGGAGGTGGAGTCTCGCTCTGTTGCCAGGCTGGAGTGCAGTGGCGAAA
TCCCTGCTCACTGCAGCCTCCGCTTCCCTGGTTCAAGCGATTCTCTTGCCCTCAGCTTCCCCA
GTAGCTGGGACCACAGGTGCCCCGCCACCACCCCAACTAATTTTTGTATTTTTTAGTAGAGAC
AGGGTTTCACCATGTTGGCCAGGCTGCTCTCAAACCCCTGACCTCAAATGATGTGCCTGCTT
CAGCCTCCCACAGTGCTGGGATTACAGGCATGGGCCACCACGCCTAGCCTCACGCTCCTTTC
TGATCTTCACTAAGAACAAAAGAAGCAGCAACTTGCAAGGGCGGCCTTTCCCACTGGTCCAT
CTGGTTTTCTCTCAGGGTCTTGCAAAATTCCTGACGAGATAAGCAGTTATGTGACCTCACG
TGCAAAGCCACCAACAGCCACTCAGAAAAGACGCACCAGCCAGAAGTGCAAGACTGCAGTC
ACTGCACGTTTTTCATCTCTAGGGACCAGAACCACCAACCCACCCCTTTCTACTTCCAAGACTTAT
TTTCACATGTGGGGAGGTTAATCTAGGAATGACTCGTTTAAGGCCTATTTTCATGATTTCTT
TGTAGCATTTGGTGCTTGACGTATTATTGTCTTTGATTCCAAATAATATGTTTCCCTTCCCT
CATTGTCTGGCGTGTCTGCGTGGACTGGTGACGTGAATCAAATCATCCACTGAAA

FIGURE 28

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA45234

><subunit 1 of 1, 453 aa, 1 stop

><MW: 49334, pI: 6.32, NX(S/T): 1

MGENDPPAVEAPFSFRSLFGLDDLKISPVAPDADAVAAQILSLLPLKFFPIIVIGIIALILA
LAIGLGIHFDCSGKYRCRSSFKCIELIARCDGVSDCKDGEDEYRCVRVGGQNAVLOVFTAAS
WKTMCSDDWKGHYANVACAQLGFPSYVSSDNLRVSSLEGQFREEFVSIDHLLPDDKV TALHH
SVYVREGCASGHVVTLQCTACGHRRGYSSRIVGGNMSLLSQWPWQASLQFQGYHLCGGSVIT
PLWIITAAHCVYDLYLPKSWTIQVGLVSLDNPAPSHLVEKIVYHSKYKPKRLGNDIALMKL
AGPLTFNEMIQPVCLPNSEENFPDGKVCWTSGWGATEDGGDASPVLNHAAPPLISNKICNHR
DVYGGIISPSMLCAGYLTGGVDSCQGDSSGGLVCQERRLWKLVGATSFGIGCAEVNKPGVYT
RVTSFLDWIHEQMERDLKT

Signal Peptide:

amino acids 1-20

Transmembrane domain:

amino acids 240-284

FIGURE 29

CCCACGCGTCCGTCCTAGTCCCCGGGCCAACTCGGACAGTTTGCTCATTTATTGCAACGGTCAAGGCTGGCTTGT
GCCAGAACGGCGCGCGCGCGCACGCACGCACACACGCGGGGAAACTTTTTTAAAAATGAAAGGCTAGAAGA
GCTCAGCGGCGGCGCGGCGCTGCGCGAGGGCTCCGGAGCTGACTCGCCGAGGCAGGAAATCCCTCCGGTCCGGA
CGCCCGCCCCGGCTCGGCGCCCGCTGGGATGGTGCAGCGCTCGCCGCCGGCCCCGAGAGCTGCTGCACTGAAG
GCCGGCGACGATGGCAGCGCGCCCGCTGCCCGTGTCCCCCGCCCGCGCCCTCCTGCTCGCCCTGGCCGGTGTCT
GCTCGCGCCCTGCGAGGCCCCGAGGGGTGAGCTTATGGAACCAAGGAAGAGCTGATGAAGTTGTCACTGCCTCTGT
TCGGAGTGGGGACCTCTGGATCCCAGTGAAGAGCTTCGACTCCAAGAATCATCCAGAAGTGCTGAATATTCGACT
ACAACGGGAAAGCAAAGAACTGATCATAAATCTGGAAAGAAATGAAGGTCTCATTGCCAGCAGTTTCACGGAAAC
CCACTATCTGCAAGACGGTACTGATGTCTCCCTCGCTCGAAATTACACGGGTCACTGTTACTACCATGGACATGT
ACGGGGATATTCTGATTAGCAGTCACTCTCAGCACGTGTTCTGGTCTCAGGGGACTTATTGTGTTTGAAGATGA
AAGCTATGTCTTAGAACCAATGAAAAGTGCAACCAACAGATACAACTCTTCCAGCGAAGAAGCTGAAAAGCGT
CCGGGGATCATGTGGATCACATCACAACACACCAACCTCGCTGCAAGAATGTGTTTCCACCACCCTCTCAGAC
ATGGGCAAGAAGGCATAAAAGAGAGACCCTCAAGGCAACTAAGTATGTGGAGCTGGTGATCGTGGCAGACAACCG
AGAGTTTTCAGAGGCAAGGAAAAGATCTGGAAAAGTTAAGCAGCGATTAAAGAGATTGCTAATCACGTTGACAA
GTTTTACAGACCACTGAACATTTCGGATCGTGTTGGTAGGCGTGGAAGTGTGGAATGACATGGACAAATGCTCTGT
AAGTCAGGACCCATTACCAGCCTCCATGAATTTCTGGACTGGAGGAAGATGAAGCTTCTACCTCGCAAAATCCCA
TGACAATGCGCAGCTTGTCACTGGGGTTTTATTTCCAAGGGACCACCATCGGCATGGCCCCAATCATGAGCATGTG
CACGGCAGACCACTGCGGGGAATTGTCTGACGACCTCAGACAATCCCTTGGTGCGAGCCGTGACCCTGGCACA
TGAGCTGGGGCCACAATTTCCGGATGAATCATGACACACTGGCAGGGGCTGTAGCTGTCAAATGGCGGTTGAGAA
AGGAGGCTGCATCATGAACGTTTCCACCGGGTACCCATTTCCCATGGTGTTCAGCAGTTGCAGCAGGAAGGACTT
GGAGACCAGCCTGGAGAAAGGAATGGGGGTGTGCCTGTTTTAACCTGCCGGAAGTCAGGGAGTCTTTCCGGGGCCA
GAAGTGTGGGAACAGATTTGTGGAAGAAGGAGAGGAGTGTGACTGTGGGGAGCCAGAGGAATGTATGAATCGCTG
CTGCAATGCCACCACCTGTACCCTGAAGCCGGACGCTGTGTGCGCACATGGGCTGTGCTGTGAAGACTGCCAGCT
GAAGCCTGCAGGAACAGCGTGCAGGGACTCCAGCAACTCCTGTGACCTCCAGAGTTCTGCACAGGGGCCAGCCC
TCACTGCCCAGCCAATGTGTACCTGCACGATGGGCACTCATGTGAGGATGTGGACGGCTACTGCTACAATGGCAT
CTGCCAGACTCACGAGCAGCAGTGTGTACGCTCTGGGGACCAGGTGCTAAACCTGCCCTGGGATCTGCTTTGA
GAGAGTCAATTCTGCAGGTGATCCTTATGGCAACTGTGGCAAAGTCTCGAAGAGTTCTTTGCCAAATGCGAGAT
GAGAGATGCTAAATGTGGAATAATCCAGTGTCAAGGAGGTGCCAGCCGGCCAGTCATTGGTACCAATGCCGTTTT
CATAGAAACAAACATCCCTCTGCAGCAAGGAGGCCGGATTCTGTGCCGGGGACCCACGTGTACTTGGGCGATGA
CATGCCGGACCCAGGGCTTGTGCTTGCAGGCACAAAGTGTGCAGATGGAAAAATCTGCCTGAATCGTCAATGTCA
AAATATTAGTGTCTTTGGGGTTTACGAGTGTGCAATGCAGTGCCACGGCAGAGGGGTGTGCAACAACAGGAAGAA
CTGCCACTGCGAGGCCCCACTGGGCACCTCCCTTCTGTGACAAGTTTGGCTTTGGAGGAAGCACAGACAGCGGCC
CATCCGGCAAGCAGAAGCAAGGCAGGAAGCTGCAGAGTCCAACAGGAGCGCGGCCAGGGCCAGGAGCCCCGTGGG
ATCGCAGGAGCATGCGTCTACTGCCTCACTGACACTCATCTGAGCCCTCCCATGACATGGAGACCGTGACCAGTG
CTGCTGCAGAGGAGGTACGCGTCCCCAAGGCCTCCTGTGACTGGCAGCATTGACTCTGTGGCTTTGCCATCGTT
TCCATGACAACAGACACAACACAGTTCTCGGGGCTCAGGAGGGGAAGTCCAGCCTACCAGGCAGCTCTGCAGAAA
CAGTGAAGGAAGGGCAGCGACTTCTGGTTGAGCTTCTGCTAAACATGGACATGCTTCAGTGCTGCTCCTGAG
AGAGTAGCAGGTTACCACTCTGGCAGGCCCCAGCCCTGCAGCAAGGAGGAAGAGGACTCAAAAGTCTGGCCTTTC
ACTGAGCCTCCACAGCAGTGGGGGAGAAGCAAGGGTTGGGCCAGTGTCCCTTTCCCCAGTGACACCTCAGCCT
TGGCAGCCCTGATGACTGGTCTCTGGCTGCAACTTAATGCTCTGATATGGCTTTTAGCATTTATTATATGAAAAT
AGCAGGGTTTTAGTTTTTAAATTTATCAGAGACCCTGCCACCCATTCCATCTCCATCCAAGCAAACCTGAATGGCAA
TGAAACAAACTGGAGAAGAAGGTAGGAGAAAGGGCGGTGAACCTCTGGCTCTTTGCTGTGGACATGCGTGACCAGC
AGTACTCAGGTTTGGGGTTTGCAGAAAGCCAGGGAACCCACAGAGTACCAACCCCTTCATTTAAACAAGTAAGAA
TGTTAAAAAGTGAAAAAATGTAAGAGCCTAACTCCATCCCCGTGGCCATTACTGCATAAAATAGAGTGCATTT
GAAAT

Variable	Unit	Value
1. α_1		0.0000
2. α_2		0.0000
3. α_3		0.0000
4. α_4		0.0000
5. α_5		0.0000
6. α_6		0.0000
7. α_7		0.0000
8. α_8		0.0000
9. α_9		0.0000
10. α_{10}		0.0000
11. α_{11}		0.0000
12. α_{12}		0.0000
13. α_{13}		0.0000
14. α_{14}		0.0000
15. α_{15}		0.0000
16. α_{16}		0.0000
17. α_{17}		0.0000
18. α_{18}		0.0000
19. α_{19}		0.0000
20. α_{20}		0.0000
21. α_{21}		0.0000
22. α_{22}		0.0000
23. α_{23}		0.0000
24. α_{24}		0.0000
25. α_{25}		0.0000
26. α_{26}		0.0000
27. α_{27}		0.0000
28. α_{28}		0.0000
29. α_{29}		0.0000
30. α_{30}		0.0000
31. α_{31}		0.0000
32. α_{32}		0.0000
33. α_{33}		0.0000
34. α_{34}		0.0000
35. α_{35}		0.0000
36. α_{36}		0.0000
37. α_{37}		0.0000
38. α_{38}		0.0000
39. α_{39}		0.0000
40. α_{40}		0.0000
41. α_{41}		0.0000
42. α_{42}		0.0000
43. α_{43}		0.0000
44. α_{44}		0.0000
45. α_{45}		0.0000
46. α_{46}		0.0000
47. α_{47}		0.0000
48. α_{48}		0.0000
49. α_{49}		0.0000
50. α_{50}		0.0000
51. α_{51}		0.0000
52. α_{52}		0.0000
53. α_{53}		0.0000
54. α_{54}		0.0000
55. α_{55}		0.0000
56. α_{56}		0.0000
57. α_{57}		0.0000
58. α_{58}		0.0000
59. α_{59}		0.0000
60. α_{60}		0.0000
61. α_{61}		0.0000
62. α_{62}		0.0000
63. α_{63}		0.0000
64. α_{64}		0.0000
65. α_{65}		0.0000
66. α_{66}		0.0000
67. α_{67}		0.0000
68. α_{68}		0.0000
69. α_{69}		0.0000
70. α_{70}		0.0000
71. α_{71}		0.0000
72. α_{72}		0.0000
73. α_{73}		0.0000
74. α_{74}		0.0000
75. α_{75}		0.0000
76. α_{76}		0.0000
77. α_{77}		0.0000
78. α_{78}		0.0000
79. α_{79}		0.0000
80. α_{80}		0.0000
81. α_{81}		0.0000
82. α_{82}		0.0000
83. α_{83}		0.0000
84. α_{84}		0.0000
85. α_{85}		0.0000
86. α_{86}		0.0000
87. α_{87}		0.0000
88. α_{88}		0.0000
89. α_{89}		0.0000
90. α_{90}		0.0000
91. α_{91}		0.0000
92. α_{92}		0.0000
93. α_{93}		0.0000
94. α_{94}		0.0000
95. α_{95}		0.0000
96. α_{96}		0.0000
97. α_{97}		0.0000

><MW: 80177, pI: 7.08, NX(S/T): 5

amino acids 1-28

[illegible][illegible]

FIGURE 32

CATCCTGCAACATGGTGAAACCACGCCTGGCTAATTTTGTGTATTTTGGTAGAGATGGGA
TTTCACCGTGTTAGCCAGGATTGTCTCAATCTGACCTCATGATCTGCCCCCTCGGCCTCCC
AAAGTGCTGGGATTACAGGCGAGTGCAACCACACCCGGCCACAACTTTTAAAGAAGTTAAT
GAAACCATACTTTTACATTTTAAATGACAGGAAAATGCTCACAATAATTGTTAACCCAAAA
TTCTGGATACAAAAGTACAATCTTTACTGTGTAAATACATGTATATGTACTATATGAAAATA
TACCAAATATCAATAATACTTATCTCTGGGTAAAAACCTCTTCTCATACCCTGTGCTAACAA
CTTTTAAACAAAAAATTTGCATCACTTTTAAAGAATCAAGAAAAATTTCTGAAGGTCATATGGG
ACAGAAAAAAAACCAAGGGAAAAATCACGCCACTTGGGAAAAAAGATTGCAAATCTGCCT
TTTTATAGATTTGTAATTAATAAGGTCCAGGCTTTCTAAGCAACTTAAATGTTTTGTTTCGA
AACAAAGTACTTGTCTGGATGTAGGAGGAAAGGGAGTGATGTCACTGCCATTATGATGCCCC
TTGAATATAAGACCCTACTTGCTATCTCCCCTGCACCAGCCAGGAGCCACCCATCCTCCAGC
ACACTGAGCAGCAAGCTGGACACACGGCACACTGATCCAAATGGGTAAGGGGATGGTGGCGA
TGCTCATTCTGGGTCTGCTACTTCTGGCGCTGCTCCTACCCGTGCAGGTTTCTTCATTTGTT
CCTTTAACCAGTATGCCGGAAGCTACTGCAGCCGAAACCACAAAGCCCTCCAACAGTGCCCT
ACAGCCTACAGCCGGTCTCCTTGTGGTCTTGCTTGCCCTTCTACATCTCTACCATTAAGAGG
CAGGTCAAGAAACAGCTACAGTTCTCCAACCCATACACTAAAACCGAATCCAAATGGTGCCT
AGAAGTTCAATGTGGCAAGGAAAAAACCAGGTCTTCATCAAATCTACTAATTTCACTCCTT
ATTAACAGAGAAACGCTTGAGAGTCTCAAACCTGGACTGGTTTAAAGAGCATCTGAAGGATTT
GACTAGATGATAAATGCCTGTACTCCAGTACTTTGGGAGGCCTAGGCCGGCGGATCACCTG
AGGTCAGGAGTTTGAGACTAACCTGGCCAAAATGGTGAAACCCCATCTGTACTAAAAATACA
AATATTGACTGGGCGTGGTGGTGAGTGCCTGTGATCCCAGCTACTCAGGTGGCTGAAGCAGG
ACAATCACTTGAACTCAGGAGGCAGAGGTTGCAGTGAGCTGAGATCGCGCTACTGCACTCTA
GCCTAGCCTGGGCAACAGAGTGAGACTTCGTCTCAAAAAAAAAAAGCCAAGTGCAGTGGCT
CACGCCTGTAATCCCGGCACCTTGGGAGGCCGAGGTGGGCGGATCACGAGGTCAGGAGATCA
AGACCATCCTGGCTAATACAGTGAAACCCTGTCTCTACTAAAAATACAAAAAATTAGCCGGG
GATGGTGGCAGGCACCTGGAGTCCCAGCTACTCGGGAGGCTGAGGCAGGAGAATAGCGTGAA
CTCAGGAGGCGGAGCTTGCAGTGAGCCGAGATTGCGCTACTGCACTCCAGCCTGGGCGACAG
CGCGAGACTCCGTCTCAAAAAAAAAAAAAAAAAAAAAA

FIGURE 33

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48309

><subunit 1 of 1, 67 aa, 1 stop

><MW: 6981, pI: 7.47, NX(S/T): 0

MGKGMVAMLILGLLLLALLLPVQVSSFVPLTSMPEATAAETTKPSNSALQPTAGLLVVLLAL
LHLYH

Signal peptide:

amino acids 15-27

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 34

GCCGCGGCGAGAGCGCGCCAGCCCCGCGCGATGCCCCGCGCGCCAGGACGCCTCCTCCCGCTGCTGGCCCCGGC
CGGCGGCCCTGACTGCGCTGCTGCTGCTGCTGCTGGGCCATGGCGGGCGGGCGCTGGGGCGCCCGGGCCCAGG
AGGCGGCGGCGGGCGGGCGGACGGGCCCCCGCGGCAGACGGCGAGGACGGACAGGACCCGCACAGCAAGCACC
TGTAACGCGCCGACATGTTACGCACGGGATCCAGAGCGCCGCGCACTTCGTCATGTTCTTCGCGCCCTGGTGTG
GACACTGCCAGCGGCTGCAGCCGACTTGGAATGACCTGGGAGACAAATACAACAGCATGGAAGATGCCAAAGTCT
ATGTGGCTAAAGTGGACTGCACGGCCCACTCCGACGTGTGCTCCGCCCAGGGGTGCGAGGATACCCCCACCTTAA
AGCTTTTCAAGCCAGGCCAAGAAGCTGTGAAGTACCAGGGTCTTCGGGACTTCCAGACACTGGAAAACCTGGATGC
TGCAGACACTGAACGAGGAGCCAGTGACACCAGAGCCGGAAGTGGAAACCGCCCACTGCCCCGAGCTCAAGCAAG
GGCTGTATGAGCTCTCAGCAAGCAACTTTGAGCTGCACGTTGCACAAGGCGACCCTTTATCAAGTTCCTTCGCTC
CGTGGTGTGGTCACTGCAAAGCCCTGGCTCCAACCTGGGAGCAGCTGGCTCTGGGCCTTGAACATTCCGAAACTG
TCAAGATTGGCAAGGTTGATTGTACACAGCACTATGAACCTCTGCTCCGAAAACAGGTTCTGGCTATCCCACTC
TTCTCTGGTTCCGAGATGGGAAAAAGGTGGATCAGTACAAGGAAAAGCGGATTTGGAGTCACTGAGGGAGTACG
TGGAGTCGCAGCTGCAGCGCACAGAGACTGGAGCGACGGAGACCGTACGCCCCCTCAGAGGCCCCGGTGTGGCAG
CTGAGCCCCGAGGCTGACAAGGGCACTGTGTTGGCACTCACTGAAAATAACTTCGATGACACCATTGCAGAAGGAA
TAACCTTCATCAAGTTTTATGCTCCATGGTGTGGTCAATTGTAAGACTCTGGCTCCTACTTGGGAGGAACTCTCTA
AAAAGGAATTCCCTGGTCTGGCGGGGTCAAGATCGCCGAAGTAGACTGCACCTGCTGAACGGAATATCTGCAGCA
AGTATTCCGTACGAGGCTACCCACGTTATTGCTTTTCCGAGGAGGGAAGAAAGTCACTGAGCACAGTGGAGGCA
GAGACCTTGACTCGTTACACCGCTTTGTCTGAGCCAAGCGAAAGACGAACCTTAGGAACACAGTTGGAGGTAC
CTCTCCTGCCAGCTCCCGCACCCCTGCGTTTAGGAGTTCACTCCACAGAGGCCACTGGGTTCCCACTGGTGGCT
GTTCAGAAAGCAGAACATACTAAGCGTGAGGTATCTTCTTTGTGTGTGTGTTTTCCAAGCCAACACACTCTACAG
ATTCTTTATTAAGTTAAGTTTTCTCTAAGTAAATGTGTAACTCACTGGTCACTGTGTAAACATTTTCAGTGGCGATA
TATCCCCCTTTGACCTTCTCTTGATGAAATTTACATGGTTTCTTTGAGACTAAAAATAGCGTTGAGGGAAATGAAA
TTGCTGGACTATTTGTGGCTCCTGAGTTGAGTGATTTTGGTGAAAGAAAGCACATCCAAAGCATAGTTTACCTGC
CCACGAGTTCTGGAAAGGTGGCCTTGTGGCAGTATTGACGTTCCCTCTGATCTTAAGGTCACAGTTGACTCAATAC
TGTGTTGGTCCGTAGCATGGAGCAGATTGAAATGCAAAAACCCACACCTCTGGAAGATACCTTCACGGCCGCTGC
TGGAGCTTCTGTTGCTGTGAATACTTCTCTCAGTGTGAGAGGTTAGCCGTGATGAAAGCAGCGTTACTTCTGACC
GTGCTGAGTAAGAGAATGCTGATGCCATACTTTATGTGTGCGATACTTGTCAAATCAGTTACTGTTCAGGGGAT
CCTTCTGTTTTCTACGGGGTGAACATGTCTTTAGTTCTCTCATGTTAACACGAAGCCAGAGCCCACATGAACTGT
TGGATGTCTTCTTAGAAAGGGTAGGCATGGAAAATTCACGAGGCTCATTCTCAGTATCTCATTAACCTCATTGA
AAGATTCCAGTTGTATTTGTACCTGGGGTGACAAGACCAGACAGGCTTTCCAGGCCCTGGGTATCCAGGGAGGC
TCTGCAGCCCTGCTGAAGGGCCCTAACTAGAGTTCTAGAGTTTCTGATTCTGTTTCTCAGTAGTCTTTTAGAGG
CTTGCTATACTTGGTCTGCTTCAAGGAGGTCGACCTTCTAATGTATGAAGAATGGGATGCATTTGATCTCAAGAC
CAAAGACAGATGTCAAGTGGGCTGCTCTGGCCCTGGTGTGCACGGCTGTGGCAGCTGTGTGATGCCAGTGTCTCTA
ACTCATGCTGTCTTGTGATTAAACACCTCTATCTCCCTTGGGAATAAGCACATACAGGCTTAAGCTCTAAGATA
GATAGGTGTTTGTCTTTTACCATCGAGCTACTTCCATAATAACCACTTTGCATCCAACACTCTTCACCCACCT
CCCATAACGCAAGGGGATGTGGATACTTGGCCCAAAGTAACCTGGTGGTAGGAATCTTAGAAACAAGACCACTTATA
CTGTCTGTCTGAGGCAGAAGATAACAGCAGCATCTCGACCAGCCTCTGCCTTAAAGGAAATCTTTATTAATCACG
TATGGTTACAGATAATTCTTTTTTTAAAAAAACCCAACTCTAGAGAAGCACAACCTGTCAAGAGTCTTGTACA
CACAACCTTCAGCTTTCATCACGAGTCTTGTATTCCAAGAAAATCAAAGTGGTACAATTTGTTTGTTTACACTAT
GATACTTTCTAAATAAACTCTTTTTTTTTTAA

FIGURE 35

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA46776

><subunit 1 of 1, 432 aa, 1 stop

><MW: 47629, pI: 5.90, NX(S/T): 0

MPARPGRLLPLLARPAALTALLLLLLGHGGGGRWGARAQEAAAAAADGPPAADGEDGQDPHS
KHLVTADMFTHGIIQSAAHFVMFFAPWCGHCQRLQPTWNDLGDKYNSMEDAKVYVAKVDCTAH
SDVCSAQGVVGYPTLKLFKPGQEAVKYQGPRDFQTLNWMQLTLNEEPVTPPEPEVEPPSAPE
LKQGLYELASNFELHVAQGDHFIKFFAPWCGHCKALAPTWEQLALGLEHSETVKIGKVDCT
QHYELCSGNQVRYPTLLWFRDGKKVDQYKGRDLESLREYVESQLQRTETGATETVTPSEA
PVLAAEPEADKGTVLALTENNFDDTIAEGITFIKFYAPWCGHCKTLAPTWEELSKKEFPGLA
GVKIAEVDCTAERNICKYSVRYPTLLLFRGGKKVSEHSGGRDLDSLHRFVLSQAKDEL

Signal sequence:

amino acids 1-32

FIGURE 36

CTTTTCTGAGGAACCACAGCAATGAATGGCTTTGCATCCTTGCTTCGAAGAAACCAATTTAT
CCTCCTGGTACTATTTCTTTTGCAAATTCAGAGTCTGGGTCTGGATATTGATAGCCGTCCTA
CCGCTGAAGTCTGTGCCACACACACAATTTACCAGGACCCAAAGGAGATGATGGTGAAAAA
GGAGATCCAGGAGAAGAGGGGAAAGCATGGCAAAGTGGGACGCATGGGGCCGAAAGGAATTAA
AGGAGAACTGGGTGATATGGGAGATCAGGGCAATATTGGCAAGACTGGGCCCATTGGGAAGA
AGGGTGACAAAGGGGAAAAAGGTTTGCTTGGAATACCTGGAGAAAAAGGCAAAGCAGGTACT
GTCTGTGATTGTGGAAGATAACCGGAAATTTGTTGGACAACCTGGATATTAGTATTGCTCGGCT
CAAGACATCTATGAAGTTTGTCAAGAATGTGATAGCAGGGATTAGGGAAACTGAAGAGAAAT
TCTACTACATCGTGCAGGAAGAGAAGAACTACAGGGAATCCCTAACCCACTGCAGGATTTCGG
GGTGGAATGCTAGCCATGCCCAGGATGAAGCTGCCAACACACTCATCGCTGACTATGTTGC
CAAGAGTGGCTTCTTTCGGGTGTTTCATTGGCGTGAATGACCTTGAAAGGGAGGGACAGTACA
TGTCACACAGACAACACTCCACTGCAGAACTATAGCAACTGGAATGAGGGGGAACCCAGCGAC
CCCTATGGTCATGAGGACTGTGTGGAGATGCTGAGCTCTGGCAGATGGAATGACACAGAGTG
CCATCTTACCATGTACTTTGTCTGTGAGTTCATCAAGAAGAAAAAGTAACTTCCCTCATCCT
ACGTATTTGCTATTTTCCTGTGACCGTCATTACAGTTATTGTTATCCATCCTTTTTTTCCTG
ATTGTACTACATTTGATCTGAGTCAACATAGCTAGAAAATGCTAAACTGAGGTATGGAGCCT
CCATCATCAAAAAAAAAAAAAAAAAA

FIGURE 37

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA50980

><subunit 1 of 1, 277 aa, 1 stop

><MW: 30645, pI: 7.47, NX(S/T): 2

MNGFASLLRRNQFILLVLFLLQIQSLGLDIDSRPTAEVCATHTISPGPKGDDGEKGDPGEEG
KHGKVGGRMGPKGIKGELGDMGDQGNIGKTGPIGKKGDKGEKGLLGIPGEKGKAGTVCDGCRY
RKFBVGQLDISIARLKTSMKFVKVNIAGIRETEEKFYIIVQEEKNYRESLTHCRIRGGMLAMP
KDEAANTLIADYVAKSGFFRVFIGVNDLEREGQYMSTDNTPLQNYSNWNEGEPSDPYGHEDC
VEMLSSGRWNDTECHLTMYFVCEFIKKKK

Signal peptide:

amino acids 1-25

100395-1040
T0420T-25E00T

FIGURE 38

GGTTCTATCGATTCTGAATTCGGCCACACTGGCCGGATCCTCTAGAGATCCCTCGACCTCGAC
CCACGCGTCCGCTGCTCTCCGCCCGTGTGGAGTGGTGGGGGCCTGGGTGGGAATGGGCGTGT
GCCAGCGCACGCGCGCTCCCTGGAAGGAGAAGTCTCAGCTAGAACGAGCGGCCCTAGGTTTT
CGGAAGGGAGGATCAGGGATGTTTGCGAGCGGCTGGAACCAGACGGTGCCGATAGAGGAAGC
GGGCTCCATGGCTGCCCTCCTGCTGCTGCCCCCTGCTGCTGTTGCTACCGCTGCTGCTGCTGA
AGCTACACCTCTGGCCGCGAGTTGCGCTGGCTTCCGGCGGACTTGGCCTTTGCGGTGCGAGCT
CTGTGCTGCAAAAGGGCTCTTCGAGCTCGCGCCCTGGCCGCGGCTGCCGCCGACCCGGAAGG
TCCCGAGGGGGGCTGCAGCCTGGCCTGGCGCCTCGCGGAACTGGCCAGCAGCGCGCGCCGCGC
ACACCTTTCTCATTCACGGCTCGCGGCGCTTTAGCTACTCAGAGGCGGAGCGCGAGAGTAAC
AGGGCTGCACGCGCCTTCCTACGTGCGCTAGGCTGGGACTGGGGACCCGACGGCGGGCGACAG
CGGCGAGGGGAGCGCTGGAGAAGGCGAGCGGGCAGCGCCGGGAGCCGGAGATGCAGCGGCCG
GAAGCGGCGCGGAGTTTGCCGGAGGGGACGGTGCCGCCAGAGGTGGAGGAGCCGCCGCCCT
CTGTCACTGGAGCAACTGTGGCGCTGCTCCTCCCCGCTGGCCCAGAGTTTCTGTGGCTCTG
GTTTCGGGCTGGCCAAGGCCGGCCTGCGCACTGCCTTTGTGCCACCCGCCCTGCGCCGGGGCC
CCCTGCTGCACTGCCTCCGCGAGCTGCGGCGCGCGCGCGCTGGTGCTGGCGCCAGAGTTTCTG
GAGTCCCTGGAGCCGGACCTGCCCGCCCTGAGAGCCATGGGGCTCCACCTGTGGGCTGCAGG
CCCAGGAACCCACCCTGCTGGAATTAGCGATTTGCTGGCTGAAGTGTCCGCTGAAGTGGATG
GGCCAGTGCCAGGATACCTCTCTTCCCCCAGAGCATAACAGACACGTGCCTGTACATCTTC
ACCTCTGGCACCACGGGCCTCCCCAAGGCTGCTCGGATCAGTCATCTGAAGATCCTGCAATG
CCAGGGCTTCTATCAGCTGTGTGGTGTCCACCAGGAAGATGTGATCTACCTCGCCCTCCAC
TCTACCACATGTCCGGTTCCCTGCTGGGCATCGTGGGCTGCATGGGCATTGGGGCCACAGTG
GTGCTGAAATCCAAGTTCTCGGCTGGTCAGTTCTGGGAAGATTGCCAGCAGCACAGGGTGAC
GGTGTTCAGTACATTGGGGAGCTGTGCCGATACCTTGTCAACCAGCCCCCGAGCAAGGCAG
AACGTGGCCATAAGGTCCGGCTGGCAGTGGGCAGCGGGCTGCGCCCAGATACCTGGGAGCGT
TTTGTGCGGCGCTTCGGGGCCCCCTGCAGGTGCTGGAGACATATGGACTGACAGAGGGCAACGT
GGCCACCATCAACTACACAGGACAGCGGGGCGCTGTGGGGCGTGCTTCCTGGCTTTACAAGC
ATATCTTCCCCTTCTCCTTGATTGCTATGATGTCAACCACAGGAGAGCCAATTCGGGACCCC
CAGGGGCACTGTATGGCCACATCTCCAGGTGAGCCAGGGCTGCTGGTGGCCCCCGTAAGCCA
GCAGTCCCCATTCTGGGCTATGCTGGCGGGCCAGAGCTGGCCCAGGGGAAGTTGCTAAAGG
ATGTCTTCCGGCCTGGGGATGTTTTCTTCAACACTGGGGACCTGCTGGTCTGCGATGACCAA
GGTTTTCTCCGCTTCCATGATCGTACTGGAGACACCTTCAGGTGGAAGGGGGAGAATGTGGC
CACAACCGAGGTGGCAGAGGTCTTCGAGGCCCTAGATTTTCTTCAGGAGGTGAACGTCTATG
GAGTCACTGTGCCAGGGCATGAAGGCAGGGCTGGAATGGCAGCCCTAGTTCTGCGTCCCCC
CACGCTTTGGACCTTATGCAGCTCTACACCCACGTGTCTGAGAACTTGCCACCTTATGCCCG
GCCCCGATTCTCTCAGGCTCCAGGAGTCTTTGGCCACCACAGAGACCTTCAAACAGCAGAAAG
TTCGGATGGCAAATGAGGGCTTCGACCCAGCACCTGTCTGACCCACTGTACGTTCTGGAC
CAGGCTGTAGGTGCCTACCTGCCCCCTCACACTGCCCCGTACAGCGCCCTCCTGGCAGGAAA
CCTTCGAATCTGAGAACTTCCACACCTGAGGCACCTGAGAGAGGAACTCTGTGGGGTGGGGG
CCGTTGCAGGTGTACTGGGCTGTGAGGGATCTTTTCTATACCAGAACTGCGGTCACTATTTT
GTAATAAATGTGGCTGGAGCTGATCCAGCTGTCTCTGACCTAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAGGGCGGCCGCGACTCTAGAGTCGACCTGCAGTAGGGATAACAGGGTAATAAGC
TTGGCCGCATGGCCCAACTTGTTTATTGCAG

FIGURE 39

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA50913

><subunit 1 of 1, 730 aa, 1 stop

><MW: 78644, pI: 7.65, NX(S/T): 2

MGVCQRTRAPWKEKSQLERAAFGFRKGGSGMFASGWNQTVPIEEAGSMAALLLLPLLLLLPL
LLLKLHLWPQLRWLPADLAFVRLCCKRALRARALAAAAADPEGPEGGCSLAWRLAELAQQ
RAAHTFLIHGSRFFSYSEAERESNRAARAFRLALGWDWGPDGGDSGEGSAGEGERAAPGAGD
AAAGSGAEFAGGDGAARGGGAAAPLSPGATVALLLPAGPEFLWLWFGGLAKAGLRTAFVPTAL
RRGPLLHCLRSCGARALVLAPEFLESLEPDLPALRAMGLHLWAAGPGTHPAGISDLLAEVSA
EVDGPVPGYLSSPQSITDTCLYIFTSGTTGLPKAARISHLKILQCQGFYQLCGVHQEDVIYL
ALPLYHMSGSLLGIVGCMGIGATVVLKSKFSAGQFWEDCQQHRVTVFQYIGELCRYLVNQPP
SKAERGHKVRRLAVGSGLRPDTWERFVRRFGPLQVLETYGLTEGNVATINYTGQRGAVGRASW
LYKHIFPFSLIRYDVTTGEPIRDPOGHCMATSPGEPGLLVAPVSQQSPFLGYAGGPELAQ GK
LLKDVFVRPGDVFFNTGDLLVCDDQGFLRFHDRTGDTFRWKGENVATTEVAEVFEALDFLQEV
NVYGVTVPGHEGRAGMAALVLRPPHALDLMQLYTHVSENLPYARPRFLRLQESLATTETFK
QQKVRMANEGFDPSTLSDPYVLDQAVGAYLPLTTARYSALLAGNLRI

Type II transmembrane domain:

amino acids 45-65

Other transmembrane domain:

amino acids 379-398

cAMP- and cGMP-dependent protein kinase phosphorylation site
starting at amino acid 136

CUB domain protein motif

amino acids 254-261

putative AMP-binding domain signature

amino acids 332-343

N-glycosylation sites

amino acids 37-40 and 483-486

FIGURE 40

CCTGTGTTAAGCTGAGGTTTCCCCTAGATCTCGTATATCCCCAACACATACCTCCACGCACA
CACATCCCCAAGAACCTCGAGCTCACACCAACAGACACACGCGCGCATAACACTCGCTCTC
GCTTGTCCATCTCCCTCCCGGGGAGCCGGCGCGCTCCACCTTTGCCGCACACTCCGGC
GAGCCGAGCCCGCAGCGCTCCAGGATTCTGCGGCTCGGAACTCGGATTGCAGCTCTGAACCC
CCATGGTGGTTTTTTTAAACACTTCTTTTCCTTCTCTTCTCGTTTTGATTGCACCGTTTTCCA
TCTGGGGGCTAGAGGAGCAAGGCAGCAGCCTTCCCAGCCAGCCCTTGTGGCTTGCCATCGT
CCATCTGGCTTATAAAAGTTTGTGAGCGCAGTCCAGAGGGCTGCGCTGCTCGTCCCCTCGG
CTGGCAGAAGGGGGTGACGCTGGGCAGCGCGGAGGAGCGCGCCGCTGCCTCTGGCGGGCTTT
CGGCTTGAGGGGCAAGGTGAAGAGCGCACCCGGCCGTGGGGTTTACCGAGCTGGATTTGTATG
TTGCACCAATGCCTTCTTGGATCGGGGCTGTGATTCTTCCCCTCTTGGGGCTGCTGCTCTCCC
TCCCCGCCGGGGCGGATGTGAAGGCTCGGAGCTGCGGAGAGGTCCGCCAGGCGTACGGTGCC
AAGGGATTGAGCCTGGCGGACATCCCCTACCAGGAGATCGCAGGGGAACACTTAAGAATCTG
TCCTCAGGAATATACATGCTGCACCACAGAAATGGAAGACAAGTTAAGCCAACAAAGCAAAC
TCGAATTTGAAAACCTTGTGGAAGAGACAAGCCATTTTGTGCGCACCACTTTTGTGTCCAGG
CATAAGAAATTTGACGAATTTTTCCGAGAGCTCCTGGAGAATGCAGAAAAAGTCACTAAATGA
TATGTTTGTACGGACCTATGGCATGCTGTACATGCAGAATTCAGAAGTCTTCCAGGACCTCT
TCACAGAGCTGAAAAGGTACTACACTGGGGGTAAATGTGAATCTGGAGGAAATGCTCAATGAC
TTTTGGGCTCGGCTCCTGGAACGGATGTTTCAGCTGATAAACCCCTCAGTATCACTTCAGTGA
AGACTACCTGGAATGTGTGAGCAAATACACTGACCAGCTCAAGCCATTTGGAGACGTGCCCC
GGAAACTGAAGATTGAGGTACCCGCGCCTTCATTGCTGCCAGGACCTTTGTCCAGGGGCTG
ACTGTGGGCAGAGAAGTTGCAAACCGAGTTTCCAAGGTCAGCCCAACCCCAAGGGTGTATCCG
TGCCCTCATGAAGATGCTGTACTGCCCATACTGTGCGGGGCTTCCCACTGTGAGGCCCTGCA
ACAACACTGTCTCAACGTCATGAAGGGCTGCTTGGCAAATCAGGCTGACCTCGACACAGAG
TGGAATCTGTTTATAGATGCAATGCTCTTGGTGGCAGAGCGACTGGAGGGGGCCATTCAACAT
TGAGTCGGTCATGGACCCGATAGATGTCAAGATTTCTGAAGCCATTATGAACATGCAAGAAA
ACAGCATGCAGGTGTCTGCAAAGGTCTTTCAGGGATGTGGTCAGCCCAAACCTGCTCCAGCC
CTCAGATCTGCCCGCTCAGCTCCTGAAAATTTTAATACACGTTTCAGGCCCTACAATCCTGA
GGAAAGACCAACAACCTGCTGCAGGCACAAGCTTGGACCGGCTGGTCACAGACATAAAAGAGA
AATTGAAGCTCTCTAAAAAGGTCTGGTCAGCATTACCCTACACTATCTGCAAGGACGAGAGC
GTGACAGCGGGCACGTCCAACGAGGAGGAATGCTGGAACGGGCACAGCAAAGCCAGATACTT
GCCTGAGATCATGAATGATGGGCTCACCAACCAGATCAACAATCCCGAGGTGGATGTGGACA
TCACTCGGCCTGACACTTTCATCAGACAGCAGATTATGGCTCTCCGTGTGATGACCAACAAA
CTAAAAAACGCCTACAATGGCAATGATGTCAATTTCCAGGACACAAGTGATGAATCCAGTGG
CTCAGGGAGTGGCAGTGGGTGCATGGATGACGTGTGTCCCACGGAGTTTGAGTTTGTACCA
CAGAGGCCCCCGCAGTGGATCCCGACCGGAGAGAGGTGGACTCTTCTGCAGCCCAGCGTGGC
CACTCCCTGCTCTCCTGGTCTCTCACCTGCATTGTCTTGGCACTGCAGAGACTGTGCAGATA
ATCTTGGGTTTTTTGGTCAGATGAAACTGCATTTTAGCTATCTGAATGGCCAACTCACTTCTT
TTCTTACACTCTTGGACAATGGACCATGCCACAAAACTTACCGTTTTTCTATGAGAAGAGAG
CAGTAATGCAATCTGCCTCCCTTTTTTGTTCCTCCAAAGAGTACCGGGTGCCAGACTGAACTG
CTTCCTCTTTCCTTCAGCTATCTGTGGGGACCTTGTTTTATTCTAGAGAGAATTCTTACTCAA
ATTTTTTCGTACCAGGAGATTTTCTTACCTTCATTTGCTTTTATGCTGCAGAAGTAAAGGAAT
CTCACGTTGTGAGGGTTTTTTTTTTTCTCATTAAAT

FIGURE 41

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA50914

><subunit 1 of 1, 555 aa, 1 stop

><MW: 62736, pI: 5.36, NX(S/T): 0

MPSWIGAVILPLLGLLLSLPAGADV KARS CGEVRQAYGAKGFS LADIPYQEIAGEHLRICPQ
EYTCCTTEMEDKLSQQSKLEFENLVEETSHFVRTTFVSRHKKFDEFFRELLEN A E KSLNDMF
VRTYGMLYMQNSEVFQDLFTELKRYYTGGNVNLEEMLNDFWARLLERMFQLINPQYHFSE DY
LECVSKYTDQLKPF GDVPRKLKIQVTRAFIAARTFVQGLTVGREVANRVSKVSPTPGCIRAL
MKMLYCPYCRGLPTVRPCNNYCLNVMKGCLANQADLDTEWNLFIDAMLLVAERLEGPFNIES
VMDPIDVKISEAIMNMQENSMQVSAKVFQCGGQPKPAPALRSARSAPENFNTRFRPYNPEER
PTTAAGTSLDRLVTDIKEKLKLSKKVWSALPYTICKDESVTAGTSNEEECWNGH SKARYLPE
IMNDGLTNQINNPEVDVDITRPDTFIRQQIMALRVMTNKLKNAYNGNDVNFQDTSDESSGSG
SGSGCMDDVCPTEFEFVTTEAPAVDPDRREVDSSAAQRGHSLLSWSLTCIVLALQRLCR

Signal peptide:

amino acids 1-23

Year	Age	Sex	Occupation	Education	Income	Health	Family	Community	Environment	Policy	Program	Impact	Outcome	Conclusion
2010	18-24	Male	Student	High School	\$10,000	Good	2 Children	Urban	Indoor Air	Low	Low	Low	Low	Low
2011	25-34	Female	Teacher	College	\$20,000	Good	1 Child	Urban	Indoor Air	Low	Low	Low	Low	Low
2012	35-44	Male	Engineer	University	\$30,000	Good	2 Children	Urban	Indoor Air	Low	Low	Low	Low	Low
2013	45-54	Female	Nurse	College	\$25,000	Good	1 Child	Urban	Indoor Air	Low	Low	Low	Low	Low
2014	55-64	Male	Manager	University	\$40,000	Good	2 Children	Urban	Indoor Air	Low	Low	Low	Low	Low
2015	65-74	Female	Retiree	High School	\$15,000	Good	1 Child	Urban	Indoor Air	Low	Low	Low	Low	Low
2016	75-84	Male	Retiree	College	\$20,000	Good	2 Children	Urban	Indoor Air	Low	Low	Low	Low	Low
2017	85-94	Female	Retiree	High School	\$10,000	Good	1 Child	Urban	Indoor Air	Low	Low	Low	Low	Low
2018	95-104	Male	Retiree	College	\$15,000	Good	2 Children	Urban	Indoor Air	Low	Low	Low	Low	Low
2019	105-114	Female	Retiree	High School	\$10,000	Good	1 Child	Urban	Indoor Air	Low	Low	Low	Low	Low
2020	115-124	Male	Retiree	College	\$15,000	Good	2 Children	Urban	Indoor Air	Low	Low	Low	Low	Low
2021	125-134	Female	Retiree	High School	\$10,000	Good	1 Child	Urban	Indoor Air	Low	Low	Low	Low	Low
2022	135-144	Male	Retiree	College	\$15,000	Good	2 Children	Urban	Indoor Air	Low	Low	Low	Low	Low
2023	145-154	Female	Retiree	High School	\$10,000	Good	1 Child	Urban	Indoor Air	Low	Low	Low	Low	Low
2024	155-164	Male	Retiree	College	\$15,000	Good	2 Children	Urban	Indoor Air	Low	Low	Low	Low	Low
2025	165-174	Female	Retiree	High School	\$10,000	Good	1 Child	Urban	Indoor Air	Low	Low	Low	Low	Low
2026	175-184	Male	Retiree	College	\$15,000	Good	2 Children	Urban	Indoor Air	Low	Low	Low	Low	Low
2027	185-194	Female	Retiree	High School	\$10,000	Good	1 Child	Urban	Indoor Air	Low	Low	Low	Low	Low
2028	195-204	Male	Retiree	College	\$15,000	Good	2 Children	Urban	Indoor Air	Low	Low	Low	Low	Low
2029	205-214	Female	Retiree	High School	\$10,000	Good	1 Child	Urban	Indoor Air	Low	Low	Low	Low	Low
2030	215-224	Male	Retiree	College	\$15,000	Good	2 Children	Urban	Indoor Air	Low	Low	Low	Low	Low
2031	225-234	Female	Retiree	High School	\$10,000	Good	1 Child	Urban	Indoor Air	Low	Low	Low	Low	Low
2032	235-244	Male	Retiree	College	\$15,000	Good	2 Children	Urban	Indoor Air	Low	Low	Low	Low	Low
2033	245-254	Female	Retiree	High School	\$10,000	Good	1 Child	Urban	Indoor Air	Low	Low	Low	Low	Low
2034	255-264	Male	Retiree	College	\$15,000	Good	2 Children	Urban	Indoor Air	Low	Low	Low	Low	Low
2035	265-274	Female	Retiree	High School	\$10,000	Good	1 Child	Urban	Indoor Air	Low	Low	Low	Low	Low
2036	275-284	Male	Retiree	College	\$15,000	Good	2 Children	Urban	Indoor Air	Low	Low	Low	Low	Low
2037	285-294	Female	Retiree	High School	\$10,000	Good	1 Child	Urban	Indoor Air	Low	Low	Low	Low	Low
2038	295-304	Male	Retiree	College	\$15,000	Good	2 Children	Urban	Indoor Air	Low	Low	Low	Low	Low
2039	305-314	Female	Retiree	High School	\$10,000	Good	1 Child	Urban	Indoor Air	Low	Low	Low	Low	Low
2040	315-324	Male	Retiree	College										

CGAGACGCGTGGGCGGACGCGTGGGCGAAAAGAACTCGGAGTGCCAAAAGCTAAATAAGTTAGCTGAGAAAAACGCACG
CAGTTTGCAGCGCCTGCGCCGGGTGCGCCAACTACGCAAAGACCAAGCGGGCTCCGCGCGGACCGGCCGCGGGG
TAGGGACCCGGCTTTGGCCCTTCAGGCTCCCCTAGCAGCGGGGAAAAGGAATTGCTGCCCGGAGTTTCTGCGGAGGT
GGAGGGAGATCAGGAAACGGCTTCTTCCCTCACTTCGCCGCCCTGGTGAGTGTGCGGGAGATTGGCAAACGCCTAGG
AAAGGACTGGGGAAAATAGCCCTGGGAAAGTGGAGAAGGTGATCAGGAGGCCGGTCCACTACGGCAGTTTATCTG
TCTGATCAGAGCCAGACGCGACGCGTCCACTTCGCAGTTCTTTCCAGGTGTGGGGACCGCAGGACAGACGGCCGA
TCCCGCCGCCCTCCGTACCAGCACTCCCAGGAGAGTCAAGCTCGCTCCCCAACGTCGAGGGCGCTCTGGCCACGA
AAAGTTCCCTGTCCACTGTGATTCTCAATTCCTTGCTTGGTFTTTTTTCTCCAGAGAACTTTTGGGTGGAGATATTA
ACTTTTTTCTTTTTTTTTTTTTCTTGGTGGAAAGCTGCTCTAGGGAGGGGGGAGGAGGAGGAGAGAAAGTGAATGTGC
TGGAGAAGAGCGAGCCCTCCTTGTTCTTCCGGAGTCCCATCCATTAAGCCATCACTTCTGGAAGATTAAAGTTGT
CGGACATGGTGAACAGCTGAGAGGAGAGGAGGATTTCTTGCCAGGTGGAGAGTCTTCACCGTCTGTTGGGTGCATG
TGTGCGCCCCGACGCGCGCGGGGCGCGTGGTTCTCCGCGTGGAGTCTCACCTGGGACCTGAGTGAATGGCTCCCCA
GGGGCTGTGCGGGGCATCCGCCCTCCGCCCTTCTCCACAGGCCCTGTGTCTGTCTGGAAAGATGCTAGCAATGGGGG
CGCTGGCAGGATTCCTGGATCCTCTGCCCTCCTCACTTATGGTTACCTGTCTGGGGCCAGGCCCTTAGAAGAGGAGG
AAGAAGGGGCCCTTACTAGCTCAAGCTGGAGAGAACTAGAGCCCAGCACAACTTCCACCTCCCAGCCCCATCTCA
TTTTTCATCCTAGCGGATGATCAGGGATTTAGAGATGTGGGTTACCACGGATCTGAGATTAAACACCTACTCTTG
ACAAGCTCGCTGCCGAAGGAGTTAACTGGAGAACTACTATGTCCAGCCTATTTGCACACCATCCAGGAGTCAGT
TTATTACTGGAAAGTATCAGATACACACCGGACTTCAACATTTCTATCATAAGACCTACCCAAACCACTGTTTAC
CTCTGGACAAATGCCACCTTACCTCAGAACTGAAGGAGGTTGGATATTC AACGCATATGGTCGAAAAATGGCACT
TGGGFTTTTAAACAGAAAAGAAATGCATGCCACCAGAAAGAGGATTTGATACCTTTTTTTGGTTCCCTTTTGGGAAGTG
GGGATTACTATACACACTACAAATGTGACAGTCTCTGGGATGTGTGGCTATGACTTGTATGAAAACGACAATGCTG
CCTGGGACTATGACAAATGGCATATACCTCCACACAGATGTACACTCAGAGAGTACAGCAAATCTTAGCTTCCCAT
ACCCACAAAAGCCTATATTTTTTATATACCTGCCCTATCAAGCTGTTTCAATTCACCACCTGCAAGCTCCTGGCAGGTATT
TCGAACACTACCGATCCATTATCAACATAAACAGGAGAAGATATGCTGCCATGCTTTCTTCTGCTTAGATGAAGCAA
TCAACAACGTGACATTGGCTCTAAAGACTTATGGTTTCTATAACAACAGCATTATCATTTACTCTTCAGATAATG
GTGGCCAGCCTACGGCAGGAGGGAGTAACTGGCCCTCTCAGAGGTAGCAAAGGAACATATTGGGAAGGAGGGATCC
GGGCTGTAGGCTTTGTGCATAGCCCACTTCTGAAAAACAAGGGAACAGTGTGTAAGGAACCTTGTGCACATCACTG
ACTGGTACCCCACTCTCATTTTCACTGGCTGAAGGACAGATTGATGAGGACATTCAACTAGATGGCTATGATATCT
GGGAGACCATAAGTGAGGGTCTTCGCTCACCCCGAGTAGATATTTTGCATAACATTGACCCCTATACACCAAGGC
AAAAAATGGCTCCTGGGCAGCAGGCCTATGGGATCTGGAACTGCAATCCAGTCAGCCATCAGAGTGCAGCACTG
GAAATTGCTTACAGGAAATCCTGGCTACAGCGACTGGGTCCCCCTCAGTCTTTTACAGCAACCTGGGACCGAACC
GTGGCACAATGAACGGATCACCTTGTCAACTGGCAAAAGTGTATGGCTTTTCAACATCACAGCCGACCCATATGA
GAGGGTGGACCTATCTAACAGGTATCCAGGAATCGTGAAGAAGCTCCTACGAGGCTCTCACAGTTCAACAAAAC
TGCAGTGCCGGTCAGGTATCCCCCAGAACCCAGAAAGTAAACCTAGGCTCAATGGAGGGGTCTGGGGACCATG
GTATAAAGAGGAAACCAAGAAAAAGAACCAAGCAAAAAATCAGGCTGAGAAAAAGCAAAAGAAAAAGCAAAAAA
GAAGAAGAAAACAGCAGAAAGCAGTCTCAGGTAAACCAGCAAATTTGGCTCGATAATATCGCTGGCCTAAGCGTCA
GGCTTGTTTTTCATGCTGTGCCACTCCAGAGACTTCTGCCACCTGGCCGCCCACTGAAAACTGTCTTGTCTCAGTG
CCAAGGTGCTACTCTTGCAAGCCACACTTAGAGAGAGTGGAGATGTTTATTTCTCTCGCTCCTTTAGAAAACGTG
GTGAGTCTGAGTTCCTACTGCTGTGCTTCACTCAACTGACCAAACTGCTTTGAATTATAGGAGGAGAAACAATA
ACCTACCATCCGCAAGCATGCTAAATTTGATGGAAAGTTACAGGGTAGCATGATTAAAACTACCTTTTGATAAAATTAC

FIGURE 42B

AGTCAAAGATTGTGTCACCTCAAAGGCCTGAAGAATATATTTTCTTGGTGAATTTTGTATGTCTGTCATATGA
CACTTGGGTTTTTTAATTAATTCTATTTTATATATATAAATATATGTTTCTTTTCTGTGAAAAGCTGTTTTTCT
CACATGTGAACAGCTTGCACCTCATTTTACCATGCGTGAGGGAATGGCAAATAAGAATGTTTGAGCACACTGCCC
ACAATGAATGTAACATTTTTCTAAACACTTTACTAGAAGAACATTTTCAGTATAAAAAACCTAATTTATTTTTACA
GAAAAATATTTTGTGTTTTTATAAAAAAGTTATGCAAATGACTTTTATTTTTATTTCTGCATACCATTAGAAGA
ATTTTATTTTCATTTCTTCAAATTATCAAGCACTGTAATACTATAAAATTAATGTAATACTGTGTGAATTCAGACTA
TAAAAACATCATTCAGAAAACTTTATAATCGTCATTGTTCAATCAAGATTTTGAATGTAATAAGATGAATATAT
ATTACTTGGAATTCATGTTTGTGTCAGAGTTGAGACAACTTTATTGTTTCTATCATAAACTATTTATGTATCTT
AATTATTAATGATTTACTTTTATGGCACTAGAAAAATTTACTGTGGCTTTTCTGATCTAAGTTCTAGCTAAAT
GTATCATTTGGTCCATAAAAAATAAAAACTTTACTAATAGGCAATTGAAGGAATGGTTTGCTAACAACACAGTAA
TATAATATGATTTTACAGATAGATGCTTCCCTTGGCTATGACATGGAGAAAGATTTTCCATAATAATACTAA
TATTTATATTAGGTTGGTGCAAACTAGTTGCGGTTTTTCCCATTAAGTAATAACCTTACTCTTATACAAAGT
GGACACTGTGGGGAGATACAGAGAAATGGAAGATACGGATCCTGCCTGGAGTAGGTAACCTTGCTTGGAACCCC
ACATGCAAACGTCATGAGGAGAATTAAAGGAGTATTATCAGTAATGAAGTTTATCATGGGTTCATCAATGAGCATA
GATTGGTGTGGATCCTGTAGACCCTGGTGTTTTCTTTGAAGTGCCCTCTCCTAATGCAGAGGCCTTGAAGCTTAC
AGTATACACTTGAAAAGTCACAGATAGCTAGAATTATGATCTTTGAAGTTATAACTGTGATCTGAAAATGTGTGT
GGTGGTATGACAGCATACCATTAAATACATTTACATCACAGCTCAAAGGACTGTGATATAATCCATTTATATCAC
AACTCAAAGGACTGTGATATAATCCATTTATATCACAGCTCACAGTTTCTGAAAATGTATAAAAGAATCTATAAT
CTAGTACTGAAATTACTAAATTTGGGTAAGATGATTTAAATGATTTTAAATTTTAAATTTTCTAGAATATAT
GGCTCCATTTTATTTTATAGTGTAAGTTGTATTTCTTAAAGTTTGTGTTTTGTGCGACAGTATCTTTTAAATGAG
TCTTAAAAATAAAGGCATATTGTTTCATGTTTAAA
AAA

[illegible]

```
><subunit 1 of 1, 515 aa, 1 stop
```

MAPRGCAGHPPPPSPQACVCPGKMLAMGALAGFWILCLLTYGYLSWGQALEEEEEEGALLAQAG
GEKLEPSTTSTSQPHLIFILADDQGFRDVGYPHYGSEIKTPTLDKLAAEGVKLENYYVQPICTP
SRSQFITGKYQIHTGLQHSIIIRPTQPNCLPLDNATLPQKLKEVGYSTHMGVKGWHLGFNRKEC
MPTRRGFDTFFGSLLGSGDYIYTHYKCDSPGMCYDLYENDNAAWDYDNGIYSTQMYTQRVQQ
ILASHNPTKPIFLYTAYQAVHSPLQAPGRYFEHYRSIININRRRYAAMLSCLDEAINNVTLA
LKTYGFYNNIIIIYSSDNGGQPTAGGSNWPLRGSKGTYWEGGIRAVGFVHSPLLKNKGTVCK
ELVHITDWYPTLISLAEGQIDEDIQLDGYDIWETISEGLRSPRVDILHNIDPYTPRQKMAPG
QQAMGSGTLQSSQPSECSTGNCLQEILATATGSPLSLSATWDRGTGTMNGSPCQLAKVYGFS
TSQPTHMRGWTYLTGIOES

Signal Peptide:

Sulfatases signature 1.

amino acids 120-132

amino acids 168-177

amino acids 163-169

amino acids 157-160, 306-309 and 318-321

Year	Age	Sex	Location	Case No.	Ref.
1991	10	M	Chennai	1	[1]
1992	15	F	Chennai	2	[2]
1993	20	M	Chennai	3	[3]
1994	25	F	Chennai	4	[4]
1995	30	M	Chennai	5	[5]
1996	35	F	Chennai	6	[6]
1997	40	M	Chennai	7	[7]
1998	45	F	Chennai	8	[8]
1999	50	M	Chennai	9	[9]
2000	55	F	Chennai	10	[10]
2001	60	M	Chennai	11	[11]
2002	65	F	Chennai	12	[12]
2003	70	M	Chennai	13	[13]
2004	75	F	Chennai	14	[14]
2005	80	M	Chennai	15	[15]
2006	85	F	Chennai	16	[16]
2007	90	M	Chennai	17	[17]
2008	95	F	Chennai	18	[18]
2009	100	M	Chennai	19	[19]
2010	105	F	Chennai	20	[20]
2011	110	M	Chennai	21	[21]
2012	115	F	Chennai	22	[22]
2013	120	M	Chennai	23	[23]
2014	125	F	Chennai	24	[24]
2015	130	M	Chennai	25	[25]
2016	135	F	Chennai	26	[26]
2017	140	M	Chennai	27	[27]
2018	145	F	Chennai	28	[28]
2019	150	M	Chennai	29	[29]
2020	155	F	Chennai	30	[30]

CGGACGCGTGGGTGCGAGTGGAGCGGAGGACCCGAGCGGCTGAGGAGAGAGGAGGCGGCGGC
TTAGCTGCTACGGGGTCCGGCCGGCGCCCTCCCGAGGGGGGCTCAGGAGGAGGAAGGAGGAC
CCGTGCGAGA**ATG**CCTCTGCCCTGGAGCCTTGCGCTCCCGCTGCTGCTCTCCTGGGTGGCAG
GTGGT**TT**TCGGGAACGCGGCCAGTGCAAGGCATCACGGGT**TG**TTAGCATCGGCACGTCAGCCT
GGGGT**CT**GTCACTATGGAAC**TAA**CTGGC**CT**GTCTGCTACGGCTGGAGAAGAAACAGCAAGGG
AGTCTGTGAAGCTACATGCGAAC**CT**GGATGTAA**GT**TTGGTGAGTGCCTGGGACCAAACAAAT
GCAGATGCTTTCCAGGATACACCGGGAAAACCTGCAGTCAAGATGTGAATGAGTGTGGAATG
AAACCCCGGCCATGCCAACACAGATGTGTGAATACACACGGAAGCTACAAGTGCTTTTGCCT
CAGTGGCCACATGCTCATGCCAGATGCTACGTGTGTGA**ACT**CTAGGACATGTGCCATGATAA
ACTGTCAGTACAGCTGTGAAGACACAGAAGAAGGGCCACAGTGCCTGTGTCCATCCTCAGGA
CTCCGCCTGGCCCCAAATGGAAGAGACTGTCTAGATATTGATGAATGTGCCTCTGGTAAAGT
CATCTGTCCCTACAATCGAAGATGTGTGAACACATTTGGAAGCTACTACTGCAAATGTCACA
TTGGT**TT**TCGAACTGCAATATATCAGTGGACGATATGACTGTATAGATATAAATGAATGTACT
ATGGATAGCCATACGTGCAGCCACCATGCCAATTGCTTCAATACCCAAGGGTCTTCAAGTG
TAAATGCAAGCAGGGATATAAAGGCAATGGACTTCGGTGT**TT**CTGCTATCCCTGAAAAT**TT**CTG
TGAAGGAAGTCTCAGAGCACCTGGTACCATCAAAGACAGAATCAAGAAGTTGCTTGCTCAC
AAAAACAGCATGAAAAAGAAGGCAAAAAT**TAAAA**ATGTTACCCAGAACCCACCAGGACTCC
TACCCCTAAGGTGAACTTGCGAGCCCTTCAACTATGAAGAGATAGTTTCCAGAGGCGGGAACT
CTCATGGAGGTAAAAAAGGGAATGAAGAGAA**ATGA**AAGAGGGGCTTGAGGATGAGAAAAGAG
AAGAGAAAGCCCTGAAGAATGACATAGAGGAGCGAAGCCTGCGAGGAGATGTGTTTTTCCCT
AAGGTGAATGAAGCAGGTGAATTCGGCCTGATTCTGCTCAAAGGAAAGCGCTAACTTCCAA
ACTGGAACATAAAGATTTAAATATCTCGGTTGACTGCAGCTTCAATCATGGGATCTGTGACT
GGAACAGGATAGAGAAGATGATTTTGACTGGAATCCTGCTGATCGAGATAATGCTATTGGC
TTCTATATGGCAGTTCGGCCCTTGGCAGGTCACAAGAAAGACATTGGCCGATTGAAACTTCT
CCTACCTGACCTGCAACCCCAAAGCAACTTCTGTTTGCTCTTTGATTACCGGCTGGCCGGAG
ACAAAGTCGGGAAACTTTCGAGTGTTTGTGAAAAACAGTAACAATGCCCTGGCATGGGAGAAG
ACCACGAGTGAGGATGAAAAGTGAAGACAGGGGAAAATTCAGTTGTATCAAGGAACTGATGC
TACCAAAGCATCATTTTTGAAGCAGAACGTGGCAAGGGCAAACCCGGCGAAATCGCAGTGG
ATGGCGTCTTGCTTGTTTCAGGCTTATGTCCAGATAGCCTTTTATCTGTGGATGACTGAATG
T**TA**CTATCTTTATATTTGACTTTGTATGTCAGTTCCCTGGTTTTTTTTGATATTGCATCATAG
GACCTCTGGCATTTTTAGAATTACTAGCTGAAAAATTGTAATGTACCAACAGAAATATTATTG
TAAGATGCCTTTCTTGATATAAGATATGCCAATATTTGCTTTAAATATCATATCACTGTATCT
TCTCAGTCATTTCTGAATCTTTCNCATTATATTATAAAATNTGGAAANGTCAGTTTATCTC
CCCTCCTCNGTATATCTGATTTGTATANGTANGTTGATGNGCTTCTCTCTACAACATTTCTA
GAAAAATAGAAAAAAAAGCACAGAGAAATGTTTAACTGTTTGACTCTTATGATACTTCTTGG
AACTATGACATCAAAGATAGACTTTTGCCTAAGTGGCTTAGCTGGGTCTTTCATAGCCAAAC
TTGTATATTTAAATTCTTTGTAATAATAA

FIGURE 45

MPLPWSLALPLLLSWVAGGFGNAASARHHGLLASARQPGVCHYGTKLACCYGWRNRNSKGVCE
ATCEPGCKFGECVGPKNKCRCFPGYTGKTCSQDVNECGMKPRPCQHRCVNTHGSKYKCFCLSGH
MLMPDATCVNSRTCAMINCQYSCEDTEEGPQCLCPSSGLRLAPNGRDCLDIDECASGKVICP
YNRRCVNTFGSYCKCHIGFELQYISGRYDCIDINECTMDSHTCSHHANCFNTQGSFKCKCK
QGYKGNGLRCSAIPENSVKEVLRAPGTIKDRIKKLLAHKNSMKKKAKIKNVTPEPTRTPK
VNLQPFNYEEIVSRGGNSHGKKGNEEK

Signal peptide:

amino acids 1-21

EGF-like domain cysteine pattern signature.

amino acids 80-91

Calcium-binding EGF-like domains

amino acids 103-124, 230-251 and 185-206

FIGURE 46

GGGAGCTGCTGCTGTGGCTGCTGGTGTGTGCGCGCTGCTCCTGCTCTTGGTGCAGCTGCTG
CGCTTCCTGAGGGCTGACGGCGACCTGACGCTACTATGGGCCGAGTGGCAGGGACGACGCCC
AGAATGGGAGCTGACTGATATGGTGGTGTGGGTGACTGGAGCCTCGAGTGGAATTGGTGAGG
AGCTGGCTTACCAGTTGTCTAAACTAGGAGTTTCTCTTGTGCTGTCAGCCAGAAGAGTGCAT
GAGCTGGAAAGGGTGAAAAGAAGATGCCTAGAGAATGGCAATTTAAAAGAAAAAGATATACT
TGTTTTGCCCCCTTGACCTGACCGACACTGGTTCCCATGAAGCGGCTACCAAAGCTGTTCTCC
AGGAGTTTGGTAGAATCGACATTCTGGTCAACAATGGTGGAATGTCCCAGCGTTCTCTGTGC
ATGGATAACCAGCTTGGATGTCTACAGAAAGCTAATAGAGCTTAACTACTTAGGGACGGTGTC
CTTGACAAAATGTGTTCTGCCTCACATGATCGAGAGGAAGCAAGGAAAGATTGTTACTGTGA
ATAGCATCCTGGGTATCATATCTGTACCTCTTTCCATTGGATACTGTGCTAGCAAGCATGCT
CTCCGGGGTTTTTTTTAATGGCCTTCGAACAGAACTTGCCACATACCCAGGTATAATAGTTTC
TAACATTTGCCCAGGACCTGTGCAATCAAATATTGTGGAGAATTCCTTAGCTGGAGAAGTCA
CAAAGACTATAGGCAATAATGGAGACCAGTCCCACAAGATGACAACCAGTCGTTGTGTGCGG
CTGATGTTAATCAGCATGGCCAATGATTTGAAAGAAGTTTGGATCTCAGAACAACCTTTCTT
GTTAGTAACATATTTGTGGCAATACATGCCAACCTGGGCCTGGTGGATAACCAACAAGATGG
GGAAGAAAAGGATTGAGAACTTTAAGAGTGGTGTGGATGCAGACTCTTCTTATTTTAAAATC
TTTAAGACAAAACATTGACTGAAAAGAGCACCTGTACTTTTCAAGCCACTGGAGGGAGAAATG
GAAAACATGAAAACAGCAATCTTCTTATGCTTCTGAATAATCAAAGACTAATTTGTGATTTT
ACTTTTTAATAGATATGACTTTGCTTCCAACATGGAATGAAATAAAAAATAAATAATAAAAG
ATTGCCATGAATCTTGCAAAA

FIGURE 47

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA36343
><subunit 1 of 1, 289 aa, 1 stop
><MW: 32268, pI: 9.21, NX(S/T): 0
MVVVWTGASSGIGEELAYQLSKLGVSLVLSARRVHELERVKRRCLENGNLKEKDILVLPLDL
TDTGSHEAATKAVLQEFGRIDILVNNGGMSQRSLCMDTSLDVYRKLIELNYLGTVSLTKCVL
PHMIERKQGKIVTVNSILGIISVPLSIGYCASKHALRGFFNGLRTELATYPGIIVSNICPGP
VQSNIVENSLAGEVTKTIGNNGDQSHKMTTSRCVRLMLISMANDLKEVWISEQPFLVTVTLW
QYMPTWAWWITNKMGGKRIENFKSGVDADSSYFKIFKTKHD
```

Important Features:

Signal Peptide:

amino acids 1-31

Transmembrane domain:

amino acids 136-157

Tyrosine kinase phosphorylation site.

106-113 and 107-114

Homologous region to Short-chain alcohol dehydrogenase

amino acids 80-90, 131-168, 1-13 and 176-185

FIGURE 48

GCGACGTGGGCACCGCCATCAGCTGTTTCGCGCGTCTTCTCCTCCAGGTGGGGCAGGGGTTTC
GGGCTGGTGGAGCATGTGCTGGGACAGGACAGCATCCTCAATCAATCCAACAGCATATTTCGG
TTGCATCTTCTACACACTACAGCTATTGTTAGGTTGCCTGCGGACACGCTGGGCCTCTGTCC
TGATGCTGCTGAGCTCCCTGGTGTCTCTCGCTGGTTCTGTCTACCTGGCCTGGATCCTGTTT
TTCGTGCTCTATGATTTCTGCATTGTTTGTATCACCACCTATGCTATCAACGTGAGCCTGAT
GTGGCTCAGTTTCCGGAAGGTCCAAGAACCCAGGGCAAGGCTAAGAGGCACTGAGCCCTCA
ACCCAAGCCAGGCTGACCTCATCTGCTTTGCTTTGGTCTTCAAGCCGCTCAGCGTGCCTGTG
GACAGCGTGGCCCCGGCCCCCAGCCTCAGGAGGGCAACACAGTCCCTGGCGAGTGGCCCC
TGGCAGGCCAGTGTGAGGAGGCAAGGAGCCACATCTGCAGCGGCTCCCTGGTGGCAGACAC
CTGGGTCTCTCACTGCTGCCCCACTGCTTTGAAAAGGCAGCAGCAACAGAACTGAATTCCTGGT
CAGTGGTCTTGGGTTCTCTGCAGCGTGAAGGACTCAGCCCTGGGGCCGAAGAGGTGGGGGTG
GCTGCCCTGCAGTTGCCAGGGCCTATAACCACTACAGCCAGGGCTCAGACCTGGCCCTGCT
GCAGCTCGCCACCCACGACCCACACACCCCTCTGCCTGCCCCAGCCCGCCATCGCTTCC
CCTTTGGAGCCTCCTGCTGGGCCACTGGCTGGGATCAGGACACCACTGATGCTCCTGGGACC
CTACGCAATCTGCGCCTGCGTCTCATCAGTCGCCCCACATGTAAGTGTATCTACAACCAGCT
GCACCAGCGACACCTGTCCAACCCGGCCCCGGCCTGGGATGCTATGTGGGGGCCCCCAGCCTG
GGGTGCAGGGCCCCCTGTGAGGGAGATTCCGGGGGCCCTGTGCTGTGCCTCGAGCCTGACGGA
CACTGGGTTTCAGGCTGGCATCATCAGCTTTGCATCAAGCTGTGCCCAGGAGGACGCTCCTGT
GCTGCTGACCAACACAGCTGCTCACAGTTCTTGGCTGCAGGCTCGAGTTTCAGGGGGCAGCTT
TCCTGGCCAGAGCCAGAGACCCCGGAGATGAGTGATGAGGACAGCTGTGTAGCCTGTGGA
TCCTTGAGGACAGCAGGTCCCCAGGACAGGACACCTCCCCATGGCCCTGGGAGGCCAGGCT
GATGCACCAGGGACAGCTGGCCTGTGGCGGAGCCCTGGTGTGAGAGGAGGCGGTGCTAACTG
CTGCCCACTGCTTCATTGGGCGCCAGGCCCCAGAGGAATGGAGCGTAGGGCTGGGGACCAGA
CCGGAGGAGTGGGGCCTGAAGCAGCTCATCCTGCATGGAGCCTACACCCACCCTGAGGGGGG
CTACGACATGGCCCTCCTGCTGCTGGCCAGCCTGTGACACTGGGAGCCAGCCTGCGGCCCC
TCTGCCTGCCCTATCCTGACCACCACCTGCCTGATGGGGAGCGTGGCTGGGTTCTGGGACGG
GCCCCGCCAGGAGCAGGCATCAGCTCCCTCCAGACAGTGCCCGTGACCCCTCCTGGGGCCTAG
GGCCTGCAGCCGGCTGCATGCAGCTCCTGGGGGTGATGGCAGCCCTATTCTGCCGGGGATGG
TGTGTACCAGTGCTGTGGGTGAGCTGCCCAGCTGTGAGGGCCTGTCTGGGGCACCACTGGTG
CATGAGGTGAGGGGCACATGGTTCTTGGCCGGGCTGCACAGCTTCGGAGATGCTTGCCAAGG
CCCCGCCAGGCCGGCGGTCTTACC GCGCTCCCTGCCTATGAGGACTGGGTGAGCAGTTTGG
ACTGGCAGGTCTACTTCGCCGAGGAACCAAGAGCCCGAGGCTGAGCCTGGAAGCTGCCTGGCC
AACATAAGCCAACCAACCAGCTGCTGACAGGGGACCTGGCCATTCTCAGGACAAGAGAATGC
AGGCAGGCAAATGGCATTACTGCCCCCTGTCTCCCCACCCTGTCATGTGTGATTCCAGGCAC
CAGGGCAGGCCCAGAAGCCCAGCAGCTGTGGGAAGGAACCTGCCTGGGGCCACAGGTGCCCA
CTCCCCACCCTGCAGGACAGGGGTGTCTGTGGACACTCCCACACCCAACTCTGCTACCAAGC
AGGCGTCTCAGCTTTCCTCCTCCTTTACTCTTTTCAGATACAATCACGCCAGCCACGTTGTTT
TGAAAATTTCTTTTFTTGGGGGGCAGCAGTTTTCTTTTTTTAACTTAAATAAATTGTTAC
AAAATAAAA

FIGURE 49

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA40571

MLLSSLVSLAGSVYLAWILFFVLYDFCIVCITTYAINVSLMWLSFRKVQEPQGKAKRHGNTV
PGEWPWQASVRRQGAHICSGSLVADTWVLTAAHCFEKAAATELNSWSVVLGSLQREGLSPGA
EEVGVAALQLPRAYNHYSQGSDLALLQLAHPTTHTPLCLPQPAHRFPFGASCWATGWDQDTS
DAPGTLRNLRLRLISRPTCNCIYNQLHQRHLSNPARGMLCGGPQPGVQGPCQGDSGGPVLC
LEPDGHWVQAGIISFASSCAQEDAPVLLTNTAAHSSWLQARVQGAAFLAQSPETPEMSDEDS
CVACGSLRTAGPQAGAPSPWPWEARLMHQGQLACGGALVSEEAVLTAAHCFIGRQAPEEWSV
GLGTRPEEWGLKQLILHGAYTHPEGGYDMALLLLAQPVTLGASLRPLCLPYPDHHLDPDGERG
WVLGRARPGAGISSLQTVPVTLTGPRACSRLLHAAPGGDGSPILPGMVCTSAVGELPSCEGLS
GAPLVHEVRGTWFLAGLHSFGDACQGPAPPAVFTALPAYEDWVSSLDWQVYFAEEPEPEAEPE
GSCLANISQPTSC

Important features:

Signal peptide:

amino acids 1-15

Homologous region to Serine proteases, trypsin family

amino acids 79-95, 343-359 and 237-247

N-glycosylation sites.

amino acids 37-40 and 564-567

Kringle domains

amino acids 79-96, 343-360 and 235-247

FIGURE 50

CGGGCCGCCCCGGCCCCATTTCGGGCCGGGCCTCGCTGCGGCCGCGACTGAGCCAGGCTGG
GCCGCGTCCCTGAGTCCCAGAGTCGGCGCGGCGCGGCAGGGGCAGCCTTCCACCACGGGGAG
CCCAGCTGTCAGCCGCCTCACAGGAAGATGCTGCGTCGGCGGGGCAGCCCTGGCATGGGTGT
GCATGTGGGTGCAGCCCTGGGAGCACTGTGGTTCTGCCTCACAGGAGCCCTGGAGGTCCAGG
TCCCTGAAGACCCAGTGGTGGCACTGGTGGGCACCGATGCCACCCTGTGCTGCTCCTTCTCC
CCTGAGCCTGGCTTCAGCCTGGCACAGCTCAACCTCATCTGGCAGCTGACAGATAACAAACA
GCTGGTGCACAGCTTTGCTGAGGGCCAGGACCAGGGCAGCGCCTATGCCAACCGCACGGCCC
TCTTCCCGACCTGCTGGCACAGGGCAACGCATCCCTGAGGCTGCAGCGCGTGCCTGTGGCG
GACGAGGGCAGCTTCACCTGCTTCGTGAGCATCCGGGATTTTCGGCAGCGCTGCCGTACGCCT
GCAGGTGGCCGCTCCCTACTCGAAGCCCAGCATGACCCTGGAGCCCAACAAGGACCTGCGGC
CAGGGGACACGGTGACCATCACGTGCTCCAGCTACCAGGGCTACCCTGAGGCTGAGGTGTTT
TGGCAGGATGGGCAGGGTGTGCCCCCTGACTGGCAACGTGACCACGTGCGAGATGGCCAACGA
GCAGGGCTTGTTTGATGTGCACAGCGTCCTGCGGGTGGTGCTGGGTGCGAATGGCACCTACA
GCTGCCTGGTGCGCAACCCCGTGCTGCAGCAGGATGCGCACRGCTCTGTCAACCATCACAGGG
CAGCCTATGACATTCCCCCCAGAGGGCCCTGTGGGTGACCGTGGGGCTGTCTGTCTGTCTCAT
TGCCTGCTGGTGGCCCTGGCTTTCGTGTGCTGGAGAAAGATCAAACAGAGCTGTGAGGAGG
AGAATGCAGGAGCTGAGGACCAGGATGGGGAGGGAGAAGGCTCCAAGACAGCCCTGCAGCCT
CTGAAACACTCTGACAGCAAAGAAGATGATGGACAAGAAATAGCCTTGACCATGAGGACCAGG
GAGCTGCTACCCCTCCCTACAGCTCCTACCCTCTGGCTGCAATGGGGCTGCACTGTGAGCCC
TGCCCCCAACAGATGCATCCTGCTCTGACAGGTGGGCTCCTTCTCAAAGGATGCGATACAC
AGACCACTGTGCAGCCTTATTTCTCCAATGGACATGATTCCCAAGTCATCCTGCTGCCTTTT
TTCTTATAGACACAATGAACAGACCACCCACAACCTTAGTTCTCTAAGTCATCCTGCCTGCT
GCCTTATTTTACAGTACATACATTTCTTAGGGACACAGTACACTGACCACATCACCACCCCTC
TTCTTCCAGTGCTGCGTGGACCATCTGGCTGCCTTTTTTCTCCAAAAGATGCAATATTCAGA
CTGACTGACCCCTGCCTTATTTACCAAAGACACGATGCATAGTACCCCGGCCTTGTTTC
TCCAATGGCCGTGATACACTAGTGATCATGTTTACGCCCTGCTTCCACCTGCATAGAATCTTT
TCTTCTCAGACAGGGACAGTGCGGCCTCAACATCTCCTGGAGTCTAGAAGCTGTTTCCTTTC
CCCTCCTTCCTCCCTGCCCCAAGTGAAGACAGGGCAGGGCCAGGAATGCTTTGGGGACACCG
AGGGGACTGCCCCCACCACCATGGTGCTATTCTGGGGCTGGGGCAGTCTTTTCCTGGC
TTGCCTCTGGCCAGCTCCTGGCCTCTGGTAGAGTGAGACTTCAGACGTTCTGATGCCTTCG
GATGTCATCTCTCCCTGCCCCAGGAATGGAAGATGTGAGGACTTCTAATTAAATGTGGGAC
TCGGAGGGATTTTGTAACCTGGGGGTATATTTTGGGGAAAATAAATGTCTTTGTAAAAAAA
AAAAAAAAAAAAA

FIGURE 51

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA41386

><subunit 1 of 1, 316 aa, 1 stop, 1 unknown

><MW: -1, pI: 4.62, NX(S/T): 4

MLRRRGSPGMGVHVGAAALGALWFCLTGALEVQVPEDPVVALVGTDATLCCSFSPPEPGFSLAQ
LNLIWQLTDTKQLVHSFAEGQDQGSAYANRTALFPDLLAQGNASLRLQVRVVADEGSFTCFV
SIRDFGSAAVSLQVAAPYSKPSMTLEPNKDLRPGDTVTITCSSYQGYPEAEVFWQDGQGVPL
TGNVTTSQMANEQGLFDVHSVLRVVLGANGTYSCLVRNPVLQQDAHXSVTITGQPMTFPPEA
LWVTVGLSVCLIALLLVALAFVCWRKIKQSC EEENAGAEDQDGE GEGSKTALQPLKHSDSKED
DGQEIA

Important features:

Signal peptide:

amino acids 1-28

Transmembrane domain:

amino acids 251-270

N-glycosylation site.

amino acids 91-94, 104-107, 189-192 and 215-218

Homologous region to Immunoglobulins and MHC

amino acids 217-234

[illegible]

TTCTGTGACCCCTTGAGAAAGAGATTGGTGGTGAATAATGTGCCACGTCTTCTAAAGAGGGGGATGC
CTGAACCTTGTCTGAAGCCCTTGTCCGTAAGCCTTGAACACTACGTTCTTAAATCTATGAAGTCTG
AGGGACCTTTTCGCTGCTTTTTGTAGGGACTTCTTTTCTTGCTTCAGCAACATGAGGCTTTTTCT
TGTGGAACGCGGTCTTGACTCTGTTCGTCACTTCTTTGATTGGGGCTTTGATCCCTGAACCA
GAAGTGAAAATTGAAGTTCCTCCAGAAGCCATTTCATCTGCCATCGCAAGACCAAAGGAGGGGA
TTTGATGTTGGTCCACTATGAAGGCTACTTAGAAAAGGACGGCTCCTTATTTCACTCCACTC
ACAAACATAACAATGGTCAGCCCATTTGGTTTACCCTGGGCATCCTGGAGGCTCTCAAAGGT
TGGGACCAGGGCTTGAAAGGAATGTGTAGGAGAGAAGAGAAAGCTCATCATTCCTCCTGCTC
TCTGGGCTATGGAAAAGAAGGAAAAGGTAAAATTCCCCCAGAAAGTACACTGATATTTAATA
TTGATCTCCTGGAGATTTCGAAATGGACCAAGATCCCATGAATCATTCCAAGAAATGGATCTT
AATGATGACTGGAAACTCTCTAAAGATGAGGTTAAAGCATATTTAAAGAAGGAGTTTGAAAA
ACATGGTGCGGTGGTGAATGAAAGTCATCATGATGCTTTGGTGGAGGATATTTTTTGATAAAG
AAGATGAAGACAAAGATGGGTTTATATCTGCCAGAGAATTTACATATAAACACGATGAGTTA
TAGAGATACATCTACCCTTTTAAATATAGCACTCATCTTTCAAGAGAGGGGCAGTCATCTTTAA
AGAACATTTTATTTTTTATACAATGTTCTTTCTTGCTTTGTTTTTTATTTTTTATATATTTTTT
CTGACTCCTATTTTAAAGAACCCCTTAGGTTTCTAAGTACCCATTTCTTTCTGATAAGTTATT
GGGAAGAAAAAGCTAATTGGTCTTTGAATAGAAGACTTCTGGACAATTTTTCACTTTCACAG
ATATGAAGCTTTGTTTTACTTTCTCACTTATAAATTTAAAATGTTGCAACTGGGAATATAACC
ACGACATGAGACCAGGTTATAGCACAAATTAGCACCCCTATATTTCTGCTTCCCTCTATTTTC
TCCAAGTTAGAGGTCAACATTTGAAAAGCCTTTTGCAATAGCCCAAGGCTTGCTATTTTCAT
GTTATAATGAAATAGTTTATGTGTAACTGGCTCTGAGTCTCTGCTTGAGGACCAGAGGAAAA
TGGTTGTTGGACCTGACTTGTTAATGGCTACTGCTTTACTAAGGAGATGTGCAATGCTGAAG
TTAGAAAACAAGGTTAATAGCCAGGCATGTTGGCTCATGCCTGTAATCCAGCACTTTGGGAG
GCTGAGGCGGGCGGATCACCTGAGGTTGGGAGTTCTGAGACCAGCCCTGACCAACACGGAGAAA
CCCTATCTCTACTAAAAATACAAAAGTAGCCCGCGCTGGTGTGATGCGCTGTAATCCAGCT
ACCCAGGAAGGCTGAGGCGGCAGAATCACTTGAACCCGAGGCCGAGGTTGCGGTAAGCCGAG
ATCACCTNCAGCCTGGACACTCTGTCTCGAAAAAAGAAAAGAACACGGTTTAATACCATATNA
ATATGTATGCATTGAGACATGCTACCTAGGACTTAAGCTGATGAAGCTTGCTTCTAGTGAT
TGGTGGCCTATTATGATAAATAGGACAAATCATTTATGTGTGAGTTTCTTTGTAATAAAATG
TATCAATATGTTATAGATGAGGTAGAAAAGTTATATTTATATTCAATATTTACTTCTTAAGGC
TAGCGGAATATCCTTCCTGGTTCTTTAATGGGTAGTCTATAGTATATTTATACTACAATAACA
TTGTATCATAAGATAAAGTAGTAAACCAGTCTACATTTTCCCATTTCTGTCTCATCAAAAAC
TGAAGTTAGCTGGGTGTGGTGGCTCATGCCTGTAATCCAGCACTTTGGGGGCCAAGGAGGG
TGGATCACTTGAGATCAGGAGTTCAAGACCAGCCTGGCCAACATGGTGAAACCTTGTCTCTA
CTAAAAAATACAAAAATTAGCCAGGCGTGGTGGTGCACACCTGTAGTCCCAGCTACTCGGGAG
GCTGAGACAGGAGATTTGCTTGAACCCGGGAGGCGGAGGTTGCAGTGAGCCAAGATTGTGCC
ACTGCACTCCAGCCTGGGTGACAGAGCAAGACTCCATCTCAAAAAAAAAAAAAAAAAAGAAGCAGA
CCTACAGCAGCTACTATTGAATAAATACCTATCCTGGATTTT

FIGURE 53

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA44194

><subunit 1 of 1, 211 aa, 1 stop

><MW: 24172, pI: 5.99, NX(S/T): 1

MRLFLWNAVLTLFVTSLLIGALIPPEVKIEVLQKPFICHRKTKGGDLMLVHYEGYLEKDGS
L
FHSTHKHNNGQPIWFTLGILEALKGWDQGLKGMCVGEKRKLIIPPALGYGKEGKGKIPPEST
LIFNIDLLEIRNGPRSHESFQEMDLNDDWKLKDEVKAYLKKEFEKHGAVVNESHHDALVED
IFDKEDEDKDGFISAREFTYKHDEL

Important features:

Signal peptide:

amino acids 1-20

N-glycosylation site.

amino acids 176-179

Casein kinase II phosphorylation site.

amino acids 143-146, 156-159, 178-181 and 200-203

Endoplasmic reticulum targeting sequence.

amino acids 208-211

FKBP-type peptidyl-prolyl cis-trans isomerase

amino acids 78-114 and 118-131

EF-hand calcium-binding domain.

amino acids 191-203, 184-203 and 140-159

S-100/ICaBP type calcium binding domain

amino acids 183-203

FIGURE 54

AATAAAGCTTCCTTAATGTTGTATATGTCTTTGAAGTACATCCGTGCATTTTTTTTTTAGCAT
CCAACCATTCCCTCCCTTGTAGTTCTCGCCCCCTCAAATCACCCCTCTCCCGTAGCCCAACCGA
CTAACATCTCAGTCTCTGAAAATGCACAGAGATGCCTGGCTACCTCGCCCTGCCTTCAGCCT
CACGGGGCTCAGTCTCTTTTTCTCTTTGGTGCCACCAGGACGGAGCATGGAGGTCACAGTAC
CTGCCACCCTCAACGTCCTCAATGGCTCTGACGCCCGCCTGCCCTGCACCTTCAACTCCTGC
TACACAGTGAACCACAAACAGTTCTCCCTGAACTGGACTTACCAGGAGTGCAACAACCTGCTC
TGAGGAGATGTTCCCTCCAGTTCCGCATGAAGATCATTAACCTGAAGCTGGAGCGGTTTCAAG
ACCGCGTGAGATTCTCAGGGAACCCAGCAAGTACGATGTGTGCGTGATGCTGAGAAACGTG
CAGCCGGAGGATGAGGGGATTTACAACCTGCTACATCATGAACCCCCCTGACCGCCACCGTG
CCATGGCAAGATCCATCTGCAGGTCCTCATGGAAGAGCCCCCTGAGCGGGACTCCACGGTG
CCGTGATTGTGGGTGCCCTCCGTGCGGGGCTTCCCTGGCTGTGGTCATCTTGGTGCTGATGGTG
GTCAAGTGTGTGAGGAGAAAAAAGAGCAGAAGCTGAGCACAGATGACCTGAAGACCGAGGA
GGAGGGCAAGACGGACGGTGAAGGCAACCCGGATGATGGCGCCAAGTAGTGGGTGGCCGGCC
CTGCAGCCTCCCGTGTCCCGTCTCCTCCCTCTCCGCCCTGTACAGTGACCCTGCCTGCTCG
CTCTTGGTGCTTCCCGTGACCTAGGACCCAGGGCCACCTGGGGCCTCCTGAACCCCCG
ACTTCGTATCTCCACCCTGCACCAAGAGTGACCCACTCTCTTCCATCCGAGAAACCTGCCA
TGCTCTGGGACGTGTGGGCCCTGGGGAGAGGAGAGAAAGGGCTCCACCTGCCAGTCCCTGG
GGGGAGGCAGGAGGCACATGTGAGGGTCCCAGAGAGAAGGGAGTGGGTGGGCAGGGGTAGA
GGAGGGGCCGCTGTACCTGCCAGTGCTTGCCTGGCAGTGGCTTCAGAGAGGACCTGGTG
GGAGGGAGGGCTTTCTGTGCTGACAGCGCTCCCTCAGGAGGGCCTTGGCCTGGCACGGCTG
TGCTCCTCCCTGCTCCCAGCCCAGAGCAGCCATCAGGCTGGAGGTGACGATGAGTTCCTGA
AACTTGAGAGGGCATGTTAAAGGGATGACTGTGCATTCCAGGGCACTGACGAAAGCCAGGG
CTGCAGGCAAAGCTGGACATGTGCCCTGGCCCAGGAGGCCATGTTGGGCCCTCGTTTCCATT
GCTAGTGGCCTCCTTGGGGCTCCTGTTGGCTCCTAATCCCTTAGGACTGTGGATGAGGCCAG
ACTGGAAGAGCAGCTCCAGGTAGGGGGCCATGTTTCCCAGCGGGGACCCACCAACAGAGGCC
AGTTTCAAAGTCAGCTGAGGGGCTGAGGGGTGGGGCTCCATGGTGAATGCAGGTTGCTGCAG
GCTCTGCCTTCTCCATGGGGTAACCACCCTCGCCTGGGCAGGGGCAGCCAAGGCTGGGAAAT
GAGGAGGCCATGCACAGGGTGGGGCAGCTTTCTTTGGGGCTTCAGTGAGAACTCTCCCAGTT
GCCCTTGGTGGGGTTTCCACCTGGCTTTTGGCTACAGAGAGGGAAGGGAAAGCCTGAGGCCG
GCATAAGGGGAGGCCTTGGAACTGAGCTGCCAATGCCAGCCCTGTCCCATCTGCGGCCACG
CTACTCGCTCCTCTCCCAACAACCTCCCTTCGTGGGGACAAAAGTGACAATTGTAGGCCAGGC
ACAGTGGCTCACGCCTGTAATCCCAGCACTTTGGGAGGCCAAGGCGGGTGGATTACCTCCAT
CTGTTTAGTAGAAATGGGGCAAAACCCATCTCTACTAAAAATACAAGAATTAGCTGGGCGTG
GTGGCGTGTGCTGTAATCCCAGCTATTTGGGAGGCTGAGGCAGGAGAATCGCTTGAGCCCG
GGAAGCAGAGGTTGCAGTGAACCTGAGATAGTGATAGTGCCACTGCAATTCAGCCTGGGTGAC
ATAGAGAGACTCCATCTCAAAAAA

FIGURE 55

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA45415

<subunit 1 of 1, 215 aa, 1 stop

<MW: 24326, pI: 6.32, NX(S/T): 4

MHRDAWLPRPAFSLTGLSLFFSLVPPGRSMEVTVPATLNVLNGSDARLPCTFNSCYTVNHKQ
FSLNWTYQECNNCSEEMFLQFRMKIINLKLERFQDRVEFSGNPSKYDVSVMRLRVQPEDEGI
YNCYIMNPPDRHRGHGKIHLQVLMEEPPERDSTVAVIVGASVGGFLAVVILVLMVVKCVRRK
KEQKLSTDDLKTEEEGKTDGEGNPDDGAK

Important features:

Signal peptide:

amino acids 1-20

Transmembrane domain:

amino acids 161-179

Immunoglobulin-like fold:

amino acids 83-127

N-glycosylation sites.

amino acids 42-45, 66-69 and 74-77

FIGURE 56

GTTGTATATGTCCTGAAGTACATCCGTGCATTTTTTTTTAGCATCCAACCATCCTCCCTTGTA
GTTCTCGCCCCCTCAAATCACCTTCTCCCTTAGCCCACCCNACTAACATCTCAGTCTCTGAA
AATGCACAGAGATGCCTGGCTACCTCGCCCTGCCTTCAGCCTCACGGGGCTCAGTCTCTTTT
TCTCTTTGGTGCCACCAGGACGGAGCATGGAGGTCCACAGTACCTGNCCACCCTCAACGTCC
TCAATGGCTCTGACGCCCCGCCTGCCCTGCCCTTCAACTCCTGCTACACAGTGAACCACAAAC
AGTTCTCCCTGAACTGGACTTACCAGGAGTGCAACAACCTGCTCTGAGGAGATGTTCTCCTCCAG
TTCCGCATGAAGATCATTAACTGAAGCTGGAGCGGTTTCAAGACCGCGTGGAGTTCTCAGG
GAACCCAGCAAGTACGATGTGTGGTGATGCTGAGAAACGTGCAGCCGGAGGATGAGGGGA
TTTACAACCTGCTACATCATGAACCCCCC

FIGURE 57

TCACGGGGCTCATCTCTTTTTCTCTTTGGTGCCCACCAGGACGGAGCATGGAGGTNCACATA
CCTGCCACCCTCAACGTCCTCAATGGCTTTGACGCCCCGCCTGCCCTGCACCTTCAACTCCNG
CTACACAGTGAACCACAAACAGTTCTCCCTGAACTGGATTTACCAGGAGTGCAACAACCTGGC
TCTGAGGAGATGTTCTCCAGTTCCCGCATGGAAGATCATTTAACCTGAAAGCTGGAAGCGG
TTTTCAAGAACCGCGTGGAAGTTTCTCAGGGAACCCAGCAAGTACGATGTGTCGGTGATGC
TGAGAAACGTGCAGCCGGAGGATGAGGGGATTTACAACCTGCTACATCATGAACCCCCC

FIGURE 58

TGCGGCGACCGTCGTACACCATGGGCCCTCCACCTCCGCCCCCTACCGTGTGGGGCTGCTCCCGGATGGCCTCCTGT
TCCTCTTGCTGCTGCTAATGCTGCTCGCGGACCCAGCGCTCCCGGCCGGACGTCACCCCCCAGTGGTGCTGGTCC
CTGGTGATTTGGGTAACCAACTGGAAGCCAAGCTGGACAAGCCGACAGTGGTGCACTACCTCTGCTCCAAGAAGA
CCGAAAGCTACTTCACAATCTGGCTGAACCTGGAACCTGCTGCTGCCTGTCATCATTGACTGCTGGATTGACAATA
TCAGGCTGGTTTACAACAAAACATCCAGGGCCACCCAGTTTCTGATGGTGTGGATGTACGTGTCCCTGGCTTTG
GGAAGACCTTCTCACTGGAGTTCTTGGACCCAGCAAAAGCAGCGTGGGTTCCTATTTCCACACCATGGTGGAGA
GCCTTGTGGGCTGGGGCTACACACGGGGTGAGGATGTCCGAGGGGCTCCCTATGACTGGCGCCGAGCCCCAAATG
AAAACGGGCCCTACTTCTTGGCCCCCTCCGCGAGATGATCGAGGAGATGTACCAGCTGTATGGGGGCCCCGTGGTG
TGGTTGCCACAGTATGGGCAACATGTACACGCTCTACTTTCTGCAGCGGCAGCCGAGGCCTGGAAGGACAAGT
ATATCCGGGCCCTTCGTGTCACTGGGTGCGCCCCGGGGGGGCGTGGCCAAGACCTGCGCGTCTGGCTTCAGGAG
ACAACAACCGGATCCAGTCATCGGGCCCCCTGAAGATCCGGGAGCAGCAGCGGTGAGTGTCTCCACCAGCTGGC
TGCTGCCCTACAACATACACATGGTCACCTGAGAAGGTGTTTCGTGCAGACACCCACAATCAACTACACACTGCGGG
ACTACCGCAAGTTCTTCCAGGACATCGGCTTTGAAGATGGCTGGCTCATGCGGCAGGACACAGAAGGGCTGGTGG
AAGCCACGATGCCACCTGGCGTGCAGCTGCCTGCTATGGTACTGGCGTCCCCACACCAGACTCCTTCTACT
ATGAGAGCTTCCCTGACCGTGACCCATAAAATCTGCTTTGGTGACGGCGATGGTACTGTGAACCTGAAGAGTGCCC
TGCAGTGCCAGGCCTGGCAGAGCCGCCAGGAGCACCAAGTGTTCGTGCAGGAGCTGCCAGGCAGCGAGCACATCG
AGATGCTGGCCAACGCCACCACCCCTGGCCTATCTGAAACGTGTGCTCCTTGGGCCCTTGACTCCTGTGCCACAGGA
CTCCTGTGGCTCGGCCGTGGACCTGCTGTTGGCCTCTGGGGCTGTATGGCCACGCGTTTGTCAAAGTTTGTGA
CTCACCATTCAAGGCCCCGAGTCTTGGACTGTGAAGCATCTGCCATGGGGAAGTGTGTTTGTATCCTTTCTCT
GTGGCAGTGAAGAAGGAAGAAATGAGAGTCTAGACTCAAGGGACACTGGATGGCAAGAATGCTGCTGATGGTGA
ACTGCTGTGACCTTAGGACTGGCTCCACAGGGTGGACTGGCTGGGCCCTGGTCCCAGTCCCTGCCTGGGGCCATG
TGTCCCCCTATTCTGTGGGCTTTTCATACCTTGCCCTACTGGGCCCTGGCCCCGAGCCTTCTATGAGGGATGTT
ACTGGGCTGTGGTCTGTACCCAGAGGTCCCAGGGATCGGCTCCTGGCCCCCTCGGGTGACCTTCCCACACACCA
GCCACAGATAGGCCTGCCACTGGTCATGGGTAGCTAGAGCTGCTGGCTTCCCTGTGGCTTAGCTGGTGGCCAGCC
TGACTGGCTTCTTGGCGAGCCTAGTAGCTCCTGCAGGCAGGGGCAGTTTGTTCGCTTCTTCGTGGTTCCCAGGC
CCTGGGACATCTCACTCCACTCCTACCTCCCTTACCACCAGGAGCATTCAGCTCTGGATTGGGCGAGCAGATGTG
CCCCCAGTCCCGCAGGCTGTGTTCCAGGGGCCCTGATTTCTCGGATGTGCTATTGGCCCCAGGACTGAAGCTGC
CTCCCTTACCCCTGGGACTGTGGTTCCAAGGATGAGAGCAGGGGTTGGAGCCATGGCCTTCTGGGAACCTATGGA
GAAAGGGAATCCAAGGAAGCAGCCAAGGCTGCTCGCAGCTTCCCTGAGCTGCACCTCTTGCTAACCCACCATCA
CACTGCCACCCTGCCCTAGGGTCTCACTAGTACCAAGTGGGTGAGCACAGGGCTGAGGATGGGGCTCCTATCCAC
CCTGGCCAGCACCCAGCTTAGTGTGGGACTAGCCAGAACTTGAATGGGACCCCTGAGAGAGCCAGGGGTCCCC
TGAGGCCCCCTAGGGGCTTCTGTCTGCCCCAGGGTGCTCCATGGATCTCCCTGTGGCAGCAGGCATGGAGAGT
CAGGGCTGCCTTCATGGCAGTAGGCTCTAAGTGGGTGACTGGCCACAGGCCGAGAAAAGGGTACAGCCTCTAGGT
GGGGTTCCCAAAGACGCCTTCAGGCTGGACTGAGCTGCTCTCCACAGGGTTTCTGTGCAGCTGGATTTTCTCTG
TTGCATACATGCCTGGCATCTGTCTCCCTTGTTCCTGAGTGGCCCCACATGGGGCTCTGAGCAGGCTGTATCTG
GATTCTGGCAATAAAAGTACTCTGGATGCTGTAAAAA

[illegible]

```
><subunit 1 of 1, 412 aa, 1 stop
```

MGLHLRPYRVGLLPDGLLFLLLLLMLLADPALPAGRHPVVLVPGDLGNQLEAKLDKPTVVH
YLCSKKTESYFTIWLNLLELLLPVIIDCWIDNIRLVYNKTSRATQFPDGDVVRVPGFGKTFSL
EFLDPSKSSVGSYFHTMVESLVGWGYTRGEDVRGAPYDWRRAPNENGPYFLALREMIEMYQ
LYGGPVVLVAHSMGNMYTLYFLQRQPQAWKDKYIRAFVSLGAPWGGVAKTLRVLASGDNNRI
PVIGPLKIREQQRSVSTSWLLPYNYTWSPEKVFVQTPTINYTLRDYRKFFQDIGFEDGWL
RQDTEGLVEATMPPGVQLHCLYGTGVPTPDSFYYESFPDRDPKICFGDGDGTVNLKSALQCQ
AWQSRQEHQVLLQELPGSEHIEMLANATTLAYLKRVLG

Signal peptide:

Potential lipid substrate binding site:

N-glycosylation sites.

amino acids 99-102, 273-276, 289-292 and 398-401

amino acids 189-201

amino acids 353-365

FIGURE 60

CGGACGCGTGGGCGGACGCGTGGGGCGGCGGCAGCGGCGGCGACGGCGACATGGAGAGCGGG
GCCTACGGCGCGGCCAAGGCGGGCGGCTCCTTCGACCTGCGGCGCTTCCTGACGCAGCCGCA
GGTGGTGGCGCGCGCCGTGTGCTTGGTCTTCGCCTTGATCGTGTTCTCCTGCATCTATGGTG
AGGGCTACAGCAATGCCCACGAGTCTAAGCAGATGTACTGCGTGTTCAACCGCAACGAGGAT
GCCTGCCGCTATGGCAGTGCCATCGGGGTGCTGGCCTTCCTGGCCTCGGCCTTCTTCTTGGT
GGTCGACGCGTATTTCCCCCAGATCAGCAACGCCACTGACCGCAAGTACCTGGTCATTGGTG
ACCTGCTCTTCTCAGCTCTCTGGACCTTCCTGTGGTTTGTGGTTTCTGCTTCCTACCAAC
CAGTGGGCAGTCACCAACCCGAAGGACGTGCTGGTGGGGGCCGACTCTGTGAGGGCAGCCAT
CACCTTCAGCTTCTTTTCCATCTTCTCCTGGGGTGTGCTGGCCTCCCTGGCCTACCAGCGCT
ACAAGGCTGGCGTGGACGACTTCATCCAGAATTACGTTGACCCCACTCCGGACCCCAACACT
GCCTACGCCTCCTACCCAGGTGCATCTGTGGACAACACTACCAACAGCCACCCTTCACCCAGAA
CGCGGAGACCACCGAGGGCTACCAGCCGCCCCCTGTGTACTGAGTGGCGGTTAGCGTGGGAA
GGGGGACAGAGAGGGCCCTCCCTCTGCCCTGGACTTTTCCATCAGCCTCCTGGAAC TGCCA
GCCCCCTCTCTTTCACCTGTTCCATCCTGTGCAGCTGACACACAGCTAAGGAGCCTCATAGCC
TGCGGGGGCTGGCAGAGCCACACCCCAAGTGCCTGTGCCCAGAGGGCTTCAGTCAGCCGCT
CACTCCTCCAGGGCACTTTTAGGAAAGGGTTTTTAGCTAGTGTTTTTCTCGCTTTTAATGA
CCTCAGCCCCGCCTGCAGTGGCTAGAAGCCAGCAGGTGCCCATGTGCTACTGACAAGTGCCT
CAGCTTCCCCCGGCCCCGGTTCAGGCCGTGGGAGCCGCTATTATCTGCGTTCTCTGCCAAAG
ACTCGTGGGGGCCATCACACCTGCCCTGTGCAGCGGAGCCGGACCAGGCTCTTGTGTCTCA
CTCAGGTTTGCTTCCCCTGTGCCCACTGCTGTATGATCTGGGGGCCACCACCCTGTGCCGGT
GGCCTCTGGGCTGCCTCCCGTGGTGTGAGGGCGGGGCTGGTGCTCATGGCACTTCCTCCTTG
CTCCACCCCTGGCAGCAGGGAAGGGCTTTGCCTGACAACACCCAGCTTTATGTAAATATTC
TGCAGTTGTTACTTAGGAAGCCTGGGGAGGGCAGGGGTGCCCCATGGCTCCAGACTCTGTC
TGTGCCGAGTGATTATAAAATCGTGGGGGAGATGCCCGCCTGGGATGCTGTTTGGAGACG
GAATAAATGTTTTCTCATTCAAAG

FIGURE 61

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48304

<subunit 1 of 1, 224 aa, 1 stop

<MW: 24810, pI: 4.75, NX(S/T): 1

MESGAYGAAKAGGSFDLRRFLTQPQVVARAVCLVFALIVFSCIYGEYGSNAHESKQMYCVFN
RNEDACRYGSAIGVLAFLASAFFLVVDAYFPQISNATDRKYLVIQDILLFSALWTFLWVVGFC
FLTQWAVTNPKDVLVGADSVRAAITFSFISFSWGLASLAYQRYKAGVDDFIQNYVDPTP
DPNTAYASYPGASVDNYQQPPFTQNAETTEGYQPPPVY

Important features:

Type II Transmembrane domain:

amino acids 1-45

Other transmembrane domains:

amino acids 74-90, 108-126 and 145-161

N-glycosylation site.

amino acids 97-100

Year	Age	Sex	Region	Country	Year	Age	Sex	Region	Country
1995	10	M	North America	USA	1995	10	M	North America	USA
1995	10	F	North America	USA	1995	10	F	North America	USA
1995	10	M	Europe	Germany	1995	10	M	Europe	Germany
1995	10	F	Europe	Germany	1995	10	F	Europe	Germany
1995	10	M	Asia	Japan	1995	10	M	Asia	Japan
1995	10	F	Asia	Japan	1995	10	F	Asia	Japan
1995	10	M	Africa	South Africa	1995	10	M	Africa	South Africa
1995	10	F	Africa	South Africa	1995	10	F	Africa	South Africa
1995	10	M	Oceania	Australia	1995	10	M	Oceania	Australia
1995	10	F	Oceania	Australia	1995	10	F	Oceania	Australia
1995	10	M	South America	Brazil	1995	10	M	South America	Brazil
1995	10	F	South America	Brazil	1995	10	F	South America	Brazil
1995	10	M	North America	Canada	1995	10	M	North America	Canada
1995	10	F	North America	Canada	1995	10	F	North America	Canada
1995	10	M	Europe	France	1995	10	M	Europe	France
1995	10	F	Europe	France	1995	10	F	Europe	France
1995	10	M	Asia	India	1995	10	M	Asia	India
1995	10	F	Asia	India	1995	10	F	Asia	India
1995	10	M	Africa	Egypt	1995	10	M	Africa	Egypt
1995	10	F	Africa	Egypt	1995	10	F	Africa	Egypt
1995	10	M	Oceania	New Zealand	1995	10	M	Oceania	New Zealand
1995	10	F	Oceania	New Zealand	1995	10	F	Oceania	New Zealand
1995	10	M	South America	Argentina	1995	10	M	South America	Argentina
1995	10	F	South America	Argentina	1995	10	F	South America	Argentina
1995	10	M	North America	Mexico	1995	10	M	North America	Mexico
1995	10	F	North America	Mexico	1995	10	F	North America	Mexico
1995	10	M	Europe	Italy	1995	10	M	Europe	Italy
1995	10	F	Europe	Italy	1995	10	F	Europe	Italy
1995	10	M	Asia	China	1995	10	M	Asia	China
1995	10	F	Asia	China	1995	10	F	Asia	China
1995	10	M	Africa	Nigeria	1995	10	M	Africa	Nigeria
1995	10	F	Africa	Nigeria	1995	10	F	Africa	Nigeria
1995	10	M	Oceania	Indonesia	1995	10	M	Oceania	Indonesia
1995	10	F	Oceania	Indonesia	1995	10	F	Oceania	Indonesia
1995	10	M	South America	Venezuela	1995	10	M	South America	Venezuela
1995	10	F	South America	Venezuela	1995	10	F	South America	Venezuela
1995	10	M	North America	Colombia	1995	10	M	North America	Colombia
1995	10	F	North America	Colombia	1995	10	F	North America	Colombia
1995	10	M	Europe	Spain	1995	10	M	Europe	Spain
1995	10	F	Europe	Spain	1995	10	F	Europe	Spain
1995	10	M	Asia	Thailand	1995	10	M	Asia	Thailand
1995	10	F	Asia	Thailand	1995	10	F	Asia	Thailand
1995	10	M	Africa	Kenya	1995	10	M	Africa	Kenya
1995	10	F	Africa	Kenya	1995	10	F	Africa	Kenya
1995	10	M	Oceania	Malaysia	1995	10	M	Oceania	Malaysia
1995	10	F	Oceania	Malaysia	1995	10	F	Oceania	Malaysia
1995	10	M	South America	Peru	1995	10	M	South America	Peru
1995	10								

GAGCCACCTACCCCTGCTCCGAGGCCAGGCCCTGCAGGGCCCTCATCGGCCAGAGGGTGATCAGTGAGCAGAAGG**ATG**
 CCCGTGGCCCGAGGCCCCCCAGGTGGCTGGCGGGCAGGGGGACGGAGGTGATGGCGAGGAAGCGAGCCAGAGGGG
 ATGTTCAAGGCCCTGTGAGGACTCCAAGAGAAAAGCCCCGGGGCTACCTCCGCCTGGTGCCCCCTGTTTGTGCTGCTG
 GCCCTGCTCGTGCTGGCTTCGGCGGGGGTGCTACTCTGGTATTTCTTAGGGTACAAGGCGGAGGTGATGGTFCAG
 CAGGTGTACTCAGGCAGTCTCGTGTACTCAATCGCCACTTCTCCAGGATCTTACCCCGCGGAATCTAGTGCC
 TTCCGCAGTGAAACCGCCAAAGCCCAAGATGCTCAAGGAGCTCATACCAAGCACCCCGCTGGGAACCTTACTAC
 AACTCCAGCTCCGTCTATTTCTTTGGGGAGGGACCCCTCACCTGCTTCTTCTGGTTCATTCTCCAAATCCCCGAG
 CACCGCCGGCTGATGCTGAGCCCCGAGGTGGTGCAGGCACCTGCTGGTGGAGGAGCTGCTGTCCACAGTCAACAG
 TCGGCTGCCGTCCCTACAGGGCCGAGTACGAAGTGGACCCCGAGGGCCTAGTGATCCTGGAAGCCAGTGTGAAA
 GACATAGCTGCATTGAATTCACAGCTGGGTTGTTACCGCTACAGCTACGTGGGCCAGGGCCAGGTCTCTCCGGCTG
 AAGGGCCCTGACCACCTGGCCCTCAGCTGCTGTGGCAGCTGCAGGGCCCCAAGGACCTCATGCTCAAATCCGG
 CTGGAGTGGACCTGCGCAGAGTGGCCGGGACCGACTGGCCATGTATGACGTGGCCGGGCCCTGGAGAAAGAGGCTC
 ATCACCTCGGTGTACGGCTGCAGCCGCCAGGAGCCCGTGGTGGAGGTTCTGGCGTGGGGGGCCATCATGGCGGT
 GTCTGGAAGAAGGGCCCTGCACAGCTACTACGACCCCTTCTGTGCTCTCCGTGCAGCCGGTGGTCTTCCAGGCCCTGT
 GAAGTGAACTGACGCTGGACAACAGGCTCGACTCCCAGGGCGTCTCAGCACCCCGTACTTCCCCAGCTACTAC
 TCGCCCCAAACCCACTGCTCTCGGCACCTCACGGTGCCCTCTCTGGACTACGGCTTGGCCCTCTGGTTTGATGCC
 TATGCTAGCTGAGGAGCGAGAAGTATGATTTGCGGTGCACCCAGGGCCAGTGGACGATCCAGAAGCAGGAGGCTGTGT
 GGCTTGCGCATCTCTGCAGCCCTACGCCGAGAGGATCCCCGTGGTGGCCACGGCCGGGATCACCATCAACTTCACC
 TCCCAGATCTCCCTCACCGGGCCCGGTGTGCGGGTGCACATATGGCTTGTTACAACCAGTCGGACCCCTGCCCTGGA
 GAGTTCTCTGTTCTGTGAATGGACTCTGTGTCCCTGCCCTGTGATGGGGTCAAGGACTGCCCCAACGGCCCTGGAT
 GAGAGAACTGCGTTTGTCAGAGCCACATTCAGTGCAAAGAGGACAGCACATGCATCTCACTGCCCAAGGTCTGT
 GATGGGCGAGCTGATTGTCTCAACGGCGCAGCGATGAAGAGCAGCTGCCAGGAAGGGGTGCCATGTGGGACATTCAAC
 TTCCAGTGTGAGGACCTGGAGCTCGGTGAAGAAGCCCCAGCCGCTGATGGCGGGCCGACTGCAGGGACCGGC
 TCGGATGAGGAGCCTGTGACTGTGGCCCTCAGGGCCCCCTCAGGCGCATTTGTTGGTGGAGCTGTGTCTCCCGAG
 GGTGAGTGGCCATGGCAGGCCAGCCTCCAGGTTTCGGGGTTCGACACATCTGTGGGGGGCCCTCATCGCTGACCGC
 TGGGTGATAACAGCTGCCCACTGCTTCCAGGAGGACAGCATGGCCCTCACGGTGCCTGTGGACCGTGTCTCTGGGC
 AAGGTGTGGCAGAACTCGCGCTGGCCTGGAGAGGTGTCTTCAAGGTGAGCCGCTGCTCCTGCACCCGTACCAC
 GAAGAGGACAGCCATGACTACGACGTGGCGCTCTGTCAGCTCGACCAACCCGGTGGTGCGCTCGGCCCGCGTGC
 CCCGTCTGCTGCCCGCGCGCTCCCACTTCTTCGAGCCCGGCCCTGCATCTGTGATTACGGCTGGGCGCCCTTG
 CCGGAGGCGCGCCCCATCAGCAACGCTTCTGCAGAAAGTGGATGTGCAGTTGATCCCCACAGGACCTGTGCGAGCGAG
 GCCTATCGCTACCAGGTGACGCCACGCGATGCTGTGTGCCGGCTACCGCAAGGGCAAGAAGGATGCCCTGTCAAGGT
 GACTCAGGTGGTCCGCTGGTGTGCAAGGCACTCAGTGGCCGCTGGTTCTTGGCGGGGCTGGTCACTGGGGCCTG
 GGCTGTGGCCGGCCTAACTACTTTCGGCGTCTACACCCGCATCACAGGTGTGATCAGCTGGATCCAGCAAGTGGTG
 ACC**TGA**GGAACTGCCCCCTGCAAAGCAGGGCCCCCTCTGGACTCAGAGAGCCAGGGCAACTGCCAAGCAGG
 GGGACAAGTATTTCCGGGGGGTGGGGGAGAGCAGGCCCTGTGGTGGCAGGAGTGGCATCTTGTCTCGTCC
 CTGATGCTGTCTCCAGTGATGGCGAGGAGTAGGAAGTGCCAGCAGCTGGGGTCAAGACGTCCTCTGAGGACC
 CAGGCCCCACCCAGCCCTTCTGCCCTCCCAATTCTCTCTCTCTCCGTCCCCCTTCTCTCACTGCTGCCAATGCAAG
 GCAGTGGCTCAGCAGCAAGAATGCTGGTTCTACATCCCGAGGAGTGTCTGAGGTGCGCCCCACTCTGTACAGAGG
 CTGTTTGGGCGAGCCTTGCCCTCCAGAGAGCAGATTCCAGCTTTCGGAAGCCCCGTGGTCTAACTTGGGATCTGGGAAT
 GGAAGGTGCTCCCATCGGAGGGGACCCCTCAGAGCCCTGGAGACTGCCAGGTGGGCGCTGCTGCCACTGTAAAGCCAA
 AAGTGGGGGAAGTCTGACTCCAGGGTCTTTGCCCAACCCCTGCTGCCACTGGGCGCTCACAGCCCAAGCCCT
 CACTGGGAGGTGAGCTCAGCTGCGCCCTTTGGAAATAAAGCTGCCCTGATCAAAAAAAAAAAAAAAAAAAAAA

[illegible]

><MW: 88846, pI: 6.41, NX(S/T): 7

amino acids 451-567

FIGURE 64

GCACCCAGGGCCAGTGGACGATCCAGAACAGGAGGCTGTGTGGCTTGCGCATCCTGCAGCCC
TACGCCGAGAGGATCCCCGTGGTGGCCACGGCCGGGATCACCATCAACTTCACCTCCCAGAT
CTCCCTCACCGGGCCCCGGTGTGCGGGTGCACTATGGCTTGTACAACCAGTCGGACCCCTGCC
CTGGAGAGTTCCTCTGTTCTGTGAATGGACTCTGTGTCCCTGCCTGTGATGGGGTCAAGGAC
TGCCCCAACGGCCTGGATGAGAGAACTGCGTTTGCAGAGCCACATTCCAGTGCAAAGAGGA
CAGCACATGCATCTCACTGCCCAAGGTCTGTGATGGGCAGCCTGATTGTCTCAACGGCAGCG
ATGAAGAGCAGTGCCAGGAAGGGGTGCCATGTGGGACATTACCTTCCAGTGTGAGGACCGG
AGCTGCGTGAAGAAGCCCAACCCGAGTGTGATGGGCGGCCCCGACTGCAGGGACGGCTCGGA
TGAGGAGCACTGTGACTGTGGCCTCCAGGGCCCCCTCCAGCCGCATTGTTGGTGGAGCTGTGT
CCTCCGAGGGTGAGTGGCCATGGCAGGCCAGCCTCCAGGTTTCGGGGTCGACACATCTGTGGG
GGGGCCCTCATCGCTGACCGCTGGGTGATAACAGCTGCCCCACTGCTTCCAGGAGGACAGCAT
GGCCTCCACGGTGCTGTGGACCGTGTTCTTGGGCAAGGTGTGGCAGAACTCGCGCTGGCCTG
GAGAGGTGTCCTTCAAGGTGAGCCGCCTGCTCCTGCACCCGTACCACGAAGAGGACAGCCAT
GACTACGACGTGGCGCTGCTGCAGCTCGACCACCCGGTGGTGCCTCGGCCGCCGTGCGCCC
CGTCTGCCTGCCCCGCGCGCTCCCACTTCTTCGAGCCCGCCTGCACTGCTGGATTACGGGCT
GGGGCGCCTTGCGCGAGGGCGGCCCCATCAGCAACGCTCTGCAGAAAGTGGATGTGCAGTTG
ATCCACAGGACCTGTGCAGCGAGGCCTATCGCTACCAGGTGACGCCACGCATGCTGTGTGC
CGGCTACCGCAAGGGCAAGAAGGATGCCTGTGAGGGTGAAGTCAAGTGGTCCGCTGGTGTGCA
AGGCACTCAGTGGCCGCTGGTTCTTGGCGGGGCTGGTCAGCTGGGGCCTGGGCTGTGGCCGG
CCTAACTACTTCGGCGTCTACACCCGCATCACAGGTGTGATCAGCTGGATCCAGCAAGTGGT
GACCTGAGGAACTGCCCCCCTGCAAAGCAGGGCCCACCTCCTGGACTCAGAGAGCCCAGGGC
AACTGCCAAGCAGGGGGACAAGTAT

FIGURE 65

GGACGAGGGCAGATCTCGTTCTGGGGCAAGCCGTTGACACTCGCTCCCTGCCACCGCCCGGG
CTCCGTGCCGCAAGTTTTTCATTTTCCACCTTCTCTGCCTCCAGTCCCCCAGCCCCTGGCCG
AGAGAAGGGTCTTACCGGCCGGGATTGCTGGAAACACCAAGAGGTGGTTTTTTGTTTTTTAAA
ACTTCTGTTTTCTTGGGAGGGGGTGTGGCGGGGCAGGATGAGCAACTCCGTTCTCTGCTCTG
TTTCTGGAGCCTCTGCTATTGCTTTGCTGCGGGGAGCCCCGTACCTTTTGGTCCAGAGGGAC
GGCTGGAAGATAAGCTCCACAAACCCAAAGCTACACAGACTGAGGTCAAACCATCTGTGAGG
TTTAACCTCCGCACCTCCAAGGACCCAGAGCATGAAGGATGCTACCTCTCCGTGCGCCACAG
CCAGCCCTTAGAAGACTGCAGTTTCAACATGACAGCTAAAACCTTTTTTCATCATTACGGAT
GGACGATGAGCGGTATCTTTGAAAACCTGGCTGCACAACTCGTGTGAGCCCTGCACACAAGA
GAGAAAGACGCCAATGTAGTTGTGGTTGACTGGCTCCCCCTGGCCCACCAGCTTTACACGGA
TGCGGTCAATAATACCAGGGTGGTGGGACACAGCATTGCCAGGATGCTCGACTGGCTGCAGG
AGAAGGACGATTTTTTCTCTCGGGAATGTCCACTTGATCGGCTACAGCCTCGGAGCGCACGTG
GCCGGGTATGCAGGCAACTTCGTGAAAGGAACGGTGGGCCGAATCACAGGTTTGGATCCTGC
CGGGCCCATGTTTGAAGGGGCCGACATCCACAAGAGGCTCTCTCCGGACGATGCAGATTTTG
TGGATGTCCTCCACACCTACACGCGTTTCCTTCGGCTTGAGCATTGGTATTGAGATGCCTGTG
GGCCACATTGACATCTACCCCAATGGGGGTGACTTCCAGCCAGGCTGTGGACTCAACGATGT
CTTGGGATCAATTGCATATGGAACAATCACAGAGGTGGTAAATGTGAGCATGAGCGAGCCG
TCCACCTCTTTGTTGACTCTCTGGTGAATCAGGACAAGCCGAGTTTTTGCCTTCCAGTGCACT
GACTCCAATCGCTTCAAAAAGGGGATCTGTCTGAGCTGCCGCAAGAACCGTTGTAATAGCAT
TGGCTACAATGCCAAGAAAATGAGGAACAAGAGGAACAGCAAAATGTACCTAAAAACCCGGG
CAGGCATGCCTTTCAGAGGTAACCTTCAGTCCCTGGAGTGTCCTTGAGGAAGGCCCTTAATA
CCTCCTTCTTAATACCATGCTGCAGAGCAGGGCACATCCTAGCCCAGGAGAAGTGGCCAGCA
CAATCCAATCAAAATCGTTGCAAAATCAGATTACACTGTGCATGTCCTAGGAAAGGGAATCTTT
ACAAAATAAACAGTGTGGACCCCTAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 66

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA49646

><subunit 1 of 1, 354 aa, 1 stop

><MW: 39362, pI: 8.35, NX(S/T): 2

MSNSVPLLCFWSLCYCFAAGSPVPFGPEGRLEDKLHKPKATQTEVKPSVRFNLRTSKDPEHE
GCYLSVGHSQPLEDCSFNMTAKTFFIIHGWTMSGIFENWLHKLVSALHTREKDANVVVVDWL
PLAHQLYTDAVNNTRVVGHSIARMLDWLQEKDDFSLGNVHLIGYSLGAHVAGYAGNFVKGTV
GRITGLDPAGPMFEGADIIHKRLSPDDADFVDVLHTYTRSFGLSIGIQMPVGHIDIYPNGGDF
QPGCGLNDVLGSIAYGTITEVVKCEHERAVHLFVDSL VNQDKPSFAFQCTDSNRFKKGICLS
CRKNRCNSIGYNAKKMRNKRNSKMYLKTRAGMPFRGNLQSLECP

Important features:

Signal peptide:

amino acids 1-16

Lipases, serine active site.

amino acids 163-172

N-glycosylation sites.

amino acids 80-83 and 136-139

FIGURE 67

CGGACGCGTGGGCGGACGCGTGGGCCCTGGGCAAGGGCCGGGGCGCCGGGCGGAGCCACCTCTTCCCCCTCCCCCGC
TTCCCTGTGCGCTCCGCTGGCTGGACGCGCTGGAGGAGTGGAGCAGCACCCGGCCGGCCCTGGGGGCTGACAGT
CGGCAAAGTTTGGCCCAAGAGGAAGTGGTCTCAAACCCGGCAGGTGGCGACCAGGCCAGACCAGGGGCGCTCG
CTGCCTGCGGGCGGGCTGTAGGCGAGGGCGCGCCCCAGTGCCGAGACCCGGGGCTTCAGGAGCCGGCCCCGGGAG
AGAAGAGTGC GGCGGGCGGACGGAGAAAACAACCTCCAAAGTTGGCGAAAGGCACCGCCCCCTACTCCC GGCTGCCG
CCGCTCCCCGCCCCCAGCCCTGGCATCCAGAGTACGGGTCGAGCCCGGGCCATGGAGCCCCCTGGGGAGGCGG
CACCAGGGAGCCTGGGCGCCCGGGGCTCCGCCGCGACCCCATCGGGTAGACCACAGAAGCTCCGGGACCCCTTCG
GCACCTCTGGACAGCCAGGATGCTGTTGGCCACCCTCCTCCTCCTCCTTGGAGGCGCTCTGGCCCATCCAG
ACCGGATTATTTTTTCAAATCATGCTTGTGAGGACCCCCCAGCAGTGCTCTTAGAAGTGACAGGGCACCTTACAGA
GGCCCCCTGGTCCGGGACAGCCGCACCTCCCCTGCCAACCTGCACCTGGCTCATCCTGGGCGACAAGGAACAGACTG
TCACCATCAGGTTCCAGAAGCTACACCTGGCCTGTGGCTCAGAGCGCTTAACCTACGCTCCCCCTCTCCAGCCAC
TGATCTCCCTGTGTGAGGCACCTCCCAGCCCTCTGCAGCTGCCCGGGGCAACGTCACCATCACTTACAGCTATG
CTGGGGCCAGAGCACCCATGGGCCAGGGCTTCCTGCTCTCCTACAGCCAAGATTGGCTGATGTGCCTGCAGGAAG
AGTTTCAGTGCCTGAACCACCGCTGTGTATCTGCTGTCCAGCGCTGTGATGGGGTTGATGCCTGTGGCGATGGCT
CTGATGAAGCAGGTTGCAGCTCAGACCCCTTCCTGGCCTGACCCCAAGACCCGTCCCCCTCCCTGCCTTGCAATG
TCACCTTGGAGGACTTCTATGGGGTCTTCTCCTCTCCTGGATATACACACCTAGCCTCAGTCTCCCACCCCCAGT
CCTGCCATTGGCTGCTGGACCCCCATGATGGCCGGCGGCTGGCCGTGCGCTTCACAGCCCTGGACTTGGGCTTTG
GAGATGCAGTGCATGTGTATGACGGCCCTGGGCCCCCTGAGAGCTCCCCGACTACTGCGTAGTCTCACCCACTTCA
GCAATGGCAAGGCTGTCACTGTGGAGACACTGTCTGGCCAGGCTGTTGTGTCTTACCACACAGTTGCTTGGAGCA
ATGGTCGTGGCTTCAATGCCACCTACCATGTGCGGGGCTATTGCTTGCCTTGGGACAGACCCTGTGGCTTAGGCT
CTGGCCTGGGAGCTGGCGAAGGCCTAGGTGAGCGCTGCTTACAGTGAGGCACAGCGCTGTGACGGCTCATGGGACT
GTGCTGACGGCACAGATGAGGAGGACTGCCCAGGCTGCCACCTGGACACTTCCCCTGTGGGGCTGCTGGCACCT
CTGGTGCCACAGCCTGCTACCTGCCTGCTGACCGTGCAACTACCAGACTTTCTGTGCTGATGGAGCAGATGAGA
GACGCTGTGGCATTGCCAGCCTGGCAATTTCCGATGCCGGGACGAGAAGTGCGTGATGAGACGTGGGTGTGCG
ATGGGCAGCCAGACTGTGCGGACGGCAGTGATGAGTGGGACTGCTCCTATGTTCTGCCCGCAAGGTCATTACAG
CTGCAGTCATTGGCAGCCTAGTGTGCGGCTGCTCCTGGTTCATCGCCCTGGGCTGCACCTGCAAGCTCTATGCCA
TTGCAACCCAGGAGTACAGCATCTTTGCCCCCTCTCCCGATGGAGGCTGAGATTGTGCAGCAGCAGGCACCCC
CTTCTACGGGCGCTCATTTGCCCAGGGTGCCATCCCACCTGTAGAAGACTTTCTACAGAGAATCCTAATGATA
ACTCAGTGCTGGGCAACCTGCGTTCTCTGCTACAGATCTTACGCCAGGATATGACTCCAGGAGGTGGCCAGGTG
CCCGCCGTGCTCAGCGGGGCGCTTGATGCGACGCTGGTACGCGCTCTCCGCCGTGGGGCTTGCTCCCTCGAA
CCAAACCCCCGGCTCGGGCTCTGAGGCCAGATCCCAGGTACACCTTCTGCTGCTCCCCCTTGAGGCCCTAGATG
GTGGCACAGGTCCAGCCCGTGAGGGCGGGGCACTGGGTGGGCAAGATGGGGAGCAGGCACCCCCACTGCCCATCA
AGGCTCCCCCTCCCATCTGCTAGCACGTCTCCAGCCCCCACTACTGTCCCTGAAGCCCCAGGGCCACTGCCCTCAC
TGCCCCCTAGAGCCATCACTATTGTCTGGAGTGGTGCAGGCCCTGCGAGGCCGCTGTTGCCAGCCTGGGGCCCC
CAGGACCAACCCGGAGCCCCCTGGACCCACACAGCAGTCTGGCCCTGGAAGATGAGGACGATGTGCTACTGG
TGCCACTGGCTGAGCCGGGGGTGTGGGTAGCTGAGGCGAGAGGATGAGCCACTGCTTACCTGAGGGGACCTGGGGG
CTCTACTGAGGCCCTCTCCCCGGGGGCTCTACTCATAGTGGCACAACCTTTTAGAGGTGGGTGAGCCTCCCCCTC
ACCACTTCCTTCCCTGTCCCTGGATTTCAGGGACTTGGTGGGCCTCCCGTTGACCCATGTAGCTGTCTATAAAGT
TAAGTGTCCCTCAGGCAGGGAGAGGGCTCACAGAGTCTCCTCTGTACGTGGCCATGGCCAGACACCCAGTCCCT
TCACCACCACCTGCTCCCCACGCCACCACCATTTGGGTGGCTGTTTTTAAAAAGTAAAGTTCTTAGAGGATCATA
GGTCTGGACACTCCATCCTTGCCAAACCTTACCCAAAAGTGGCCTTAAGCACCGGAATGCCAATTAAGTAGAGA
CCCTCCAGCCCCCAAGGGGAGGATTTGGGCAGAACCTGAGGTTTTGCCATCCACAATCCCTCCTACAGGGCCTGG
CTCACAAAAGAGTGCAACAAATGCTTCTATTCCATAGCTACGGCATTGCTCAGTAAGTTGAGGTCAAAAATAAA
GGAATCATACATCTC

FIGURE 68

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA49631

<subunit 1 of 1, 713 aa, 1 stop

<MW: 76193, pI: 5.42, NX(S/T): 4

MLLATLLLLLLGGALAHDPRIIFPNHACEDPPAVLLEVQGTQLQRPLVRDSRTSPANCTWLIL
GSKEQTVTIRFQKLHLACGSERLTLRSPLQPLISLCEAPPSPLQLPGGNVTITYSYAGARAP
MGQGFLLSYSQDWLMCLQEETFQCLNHRCVSAVQRCGDGDGSDGAGCSSDPFPGLTPRP
VPSLPCNVTLEDFYGVFSSPGYTHLASVSHPQSCHWLLDPHDGRRRLAVRFTALDLGFGDAVH
VYDGPGPPESSRLLRSLTHFSNGKAVTVETLSGQAVVSYHTVAWSNGRGFNATYHVRGYCLP
WDRPCGLGSGLGAGEGLGERCYSEAQRCDGSWDACDGTDEEDCPGCPPGHFPCGAAGTSGAT
ACYLPADRCNYQTFCADGADERRCRHCQPGNFRRCRDEKCVYETWVCDGQPCDAGSDEWDGCS
YVLPRKVITAAVIGSLVCGLLLVIALGCTCKLYAIRTQEYSIFAPLSRMEAEIVQQQAPPSY
GQLIAQGAIPPVEDFPTENPNDNSVLGNLRSLLQILRQDMTPGGGPGARRRQGRMLMRRLVR
RLRRWGLLPRTNTPARASEARSQVTPSAAPLEALDGGTGPAREGGAVGGQDGEQAPPLPIKA
PLPSASTSPAPTTVPEAPGPLPSLPLEPSLLSGVVQALRGRLLPGLPPGPTRSPPGPHTAV
LALEDEDDVLLVPLAEPGVWVAEAEDEPLLT

Important features:

Signal peptide:

amino acids 1-16

Transmembrane domain:

amino acids 442-462

LDL-receptor class A (LDLRA) domain proteins

amino acids 411-431, 152-171, 331-350 and 374-393

FIGURE 69

CGAGCTGGGCGAGAAGTAGGGGAGGGCGGTGCTCCGCCGCGGTGGCGGTTGCTATCGCTTCG
CAGAACCTACTCAGGCAGCCAGCTGAGAAGAGTTGAGGGAAAGTGCTGCTGCTGGGTCTGCA
GACGCGATGGATAACGTGCAGCCGAAAATAAAACATCGCCCCTTCTGCTTCAGTGTGAAAGG
CCACGTGAAGATGCTGCGGCTGGCACTAACTGTGACATCTATGACCTTTTTTATCATCGCAC
AAGCCCCTGAACCATATATTGTTATCACTGGATTTGAAGTCACCGTTATCTTATTTTTTCATA
CTTTTATATGTACTCAGACTTGATCGATTAATGAAGTGGTTATTTTGGCCTTTGCTTGATAT
TATCAACTCACTGGTAACAACAGTATTCATGCTCATCGTATCTGTGTTGGCACTGATACCAG
AAACCACAACATTGACAGTTGGTGGAGGGGTGTTTGCACTTGTGACAGCAGTATGCTGTCTT
GCCGACGGGGCCCTTATTTACCGGAAGCTTCTGTTCAATCCCAGCGGTCCTTACCAGAAAAA
GCCTGTGCATGAAAAAAAAAGAAGTTTTGTAATTTTATATTACTTTTTAGTTTGATACTAAGT
ATTAAACATATTTCTGTATTCTTCCAAAAAAAAAAAAAAAAAAAA

FIGURE 70

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA49645

><subunit 1 of 1, 152 aa, 1 stop

><MW: 17170, pI: 9.62, NX(S/T): 1

MDNVQPKIKHRPFCFSVKGHVKMLRLALTVTSMTFFIIAQAPEPYIVITGFEVTVILFFILL
YVLRDLRLMKWLFWPLLDIINSLVTTVFMLIVSVLALIPETTTTLTVGGGVFALVTAVCCLAD
GALIYRKLLFNPSGPYQKKPVHEKKEVL

Important features:

Potential type II transmembrane domain:

amino acids 26-42

Other potential transmembrane domain:

amino acids 44-65, 81-101 and 109-129

Leucine zipper pattern

amino acids 78-99 and 85-106

N-myristoylation site.

amino acids 110-115

Ribonucleotide reductase large subunit protein

amino acids 116-127

FIGURE 72

CAGCCCCGCGCGCCGGCCGAGTCGCTGAGCCGCGGCTGCCGGACGGGACGGGACCGGCTAGG
CTGGGCGCGCCCCCGGGCCCCGCGTGGGCATGGGCGCACTGGCCCCGGGCGCTGCTGCTGC
CTCTGCTGGCCCAGTGGCTCCTGCGCGCCGCCCGGAGCTGGCCCCCGCGCCCTTCACGCTG
CCCCCTCCGGGTGGCCGCGGCCACGAACCGCGTAGTTGCGCCACCCCGGGACCCGGGACCCCC
TGCCGAGCGCCACGCCGACGGCTTGGCGCTCGCCCTGGAGCCTGCCCTGGCGTCCCCCGCGG
GCGCCGCCAACTTCTTGGCCATGGTAGACAACCTGCAGGGGGACTCTGGCCGCGGCTACTAC
CTGGAGATGCTGATCGGGACCCCCCGCAGAAGCTACAGATTCTCGTTGACACTGGAAGCAG
TAACTTTGCCGTGGCAGGAACCCCGCACTCCTACATAGACACGTACTTTGACACAGAGAGGT
CTAGCACATACCGCTCCAAGGGCTTTGACGTACAGTGAAGTACACACAAGGAAGCTGGACG
GGCTTCGTTGGGGAAGACCTCGTCACCATCCCCAAAGGCTTCAATACTTCTTTTCTTGTCAA
CATTGCCACTATTTTTGAATCAGAGAATTTCTTTTTGCCTGGGATTAAATGGAATGGAATAC
TTGGCCTAGCTTATGCCACACTTGCCAAGCCATCAAGTTCTCTGGAGACCTTCTTCGACTCC
CTGGTGACACAAGCAAACATCCCCAACGTTTTCTCCATGCAGATGTGTGGAGCCGGCTTGCC
CGTTGCTGGATCTGGGACCAACGGAGGTAGTCTTGTCTTGGGTGGAATTGAACCAAGTTTGT
ATAAAGGAGACATCTGGTATACCCCTATTAAGGAAGAGTGGTACTACCAGATAGAAATTCTG
AAATTGGAAATTGGAGGCCAAAGCCTTAATCTGGACTGCAGAGAGTATAACGCAGACAAGGC
CATCGTGACAGTGGCACCACGCTGCTGCGCCTGCCCCAGAAGGTGTTTGATGCGGTGGTGG
AAGCTGTGGCCCGCGCATCTCTGATTCCAGAATTCTCTGATGGTTTCTGGACTGGGTCCCAG
CTGGCGTGCTGGACGAATTCGGAAACACCTTGGTCTTACTTCCCTAAAATCTCCATCTACCT
GAGAGACGAGAACTCCAGCAGGTCATTCCGTATCACAATCCTGCCTCAGCTTTACATTACAGC
CCATGATGGGGGCCGGCCTGAATTATGAATGTTACCGATTTCGGCATTTCCTCCATCCACAAAT
GCGCTGGTGATCGGTGCCACGGTGATGGAGGGCTTCTACGTCACTTCGACAGAGCCCAGAA
GAGGGTGGGCTTCGCAGCGAGCCCCGTGTGCAGAAATTGCAGGTGCTGCAGTGTCTGAAATTT
CCGGGCCTTTCTCAACAGAGGATGTAGCCAGCAACTGTGTCCCCGCTCAGTCTTTGAGCGAG
CCCATTTTGTGGATTGTGTCCTATGCGCTCATGAGCGTCTGTGGAGCCATCCTCCTTGTCTT
AATCGTCCTGCTGCTGCTGCCGTTCCGGTGTCAGCGTCGCCCCCGTGACCCTGAGGTCGTCA
ATGATGAGTCCTCTCTGGTCAGACATCGCTGGAAATGAAATAGCCAGGCCTGACCTCAAGCAA
CCATGAACTCAGCTATTAAGAAAATCACATTTCCAGGGCAGCAGCCGGGATCGATGGTGGCG
CTTTCTCCTGTGCCACCCGTCTTCAATCTCTGTTCTGCTCCCAGATGCCTTCTAGATTAC
TGTCTTTTGATTCTTGATTTTCAAGCTTTCAAATCCTCCCTACTTCCAAGAAAAATAATTAA
AAAAAAACTTCATTCTAA

FIGURE 73

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA45493

><subunit 1 of 1, 518 aa, 1 stop

><MW: 56180, pI: 5.08, NX(S/T): 2

MGALARALLLPLLAQWLLRAAPELAPAPFTLPLRVAAATNRVVAPTPGPGTPAERHADGLAL
ALEPALASPAGAAANFLAMVDNLQGDSEGRGYYLEMLIGTPPQKLQILVDTGSSNFAVAGTPHS
YIDTYFDTERSSTYRSKGFDTVVKYTQGSWTGFGEDLVTIPKGFNTSFLVNIATIFESNF
FLPGIKWNGILGLAYATLAKPSSSLETFFDSLVTQANIPNVFSMQMCGAGLPVAGSGTNGGS
LVLGGIEPSLYKGDIWYTPIKEEWYYQIEILKLEIGGQSLNLDCREYNADKAIVDSGTTLLR
LPQKVFDVAVEAVARASLIPEFSDGFWTGSQACWTNSETPWSYFPKISIIYLRDENSRSFR
ITILPQLYIQPMMGAGLNYECYRFGISPSTNALVIGATVMEGFYVIFDRAQKRVGFAASPCA
EIAGAAVSEISGPFSTEDVASNCVPAQSLSEPILWIVSYALMSVCGAILLVLLVLLLLLPFRC
QRRPRDPEVVNDESSLVRHRWK

Important features:

Signal peptide:

amino acids 1-20

Transmembrane domain:

amino acids 466-494

N-glycosylation sites.

amino acids 170-173 and 366-369

Leucine zipper pattern.

amino acids 10-31 and 197-118

Eukaryotic and viral aspartyl proteases

amino acids 109-118, 252-261 and 298-310

FIGURE 74

CGCCTCCGCCTTCGGAGGCTGACGCGCCCGGGCGCCGTTCCAGGCCTGTGCAGGGCGGATCG
GCAGCCGCCTGGCGGCGATCCAGGGCGGTGCGGGGCCTGGGCGGAGCCGGGAGGCGCGGCC
GGCATGGAGGCGCTGCTGCTGGGCGCGGGGTTGCTGCTGGGCGCTTACGTGCTTGTCTACTA
CAACCTGGTGAAGGCCCCGCCGTGCGGCGGCATGGGCAACCTGCGGGGCCGCACGGCCGTGG
TCACGGGCGCCAACAGCGGCATCGGAAAGATGACGGCGCTGGAGCTGGCGCGCCGGGGAGCG
CGCGTGGTGCTGGCCTGCCGCAGCCAGGAGCGCGGGGAGGCGGCTGCCTTCGACCTCCGCCA
GGAGAGTGGAACAATGAGGTCATCTTCATGGCCTTGGAATTGGCCAGTCTGGCCTCGGTGC
GGGCCTTTGCCACTGCCTTTCTGAGCTCTGAGCCACGGTTGGACATCCTCATCCACAATGCC
GGTATCAGTTCCTGTGGCCGGACCCGTGAGGCGTTTAACTGCTGCTTCGGGTGAACCATAT
CGGTCCCTTTCTGCTGACACATCTGCTGCTGCCTTGCTGAAGGCATGTGCCCCCTAGCCGCG
TGGTGGTGGTAGCCTCAGCTGCCCAGTCTCGGGGACGTCTTGACTTCAAACGCCTGGACCGC
CCAGTGGTGGGCTGGCGGCAGGAGCTGCGGGCATATGCTGACACTAAGCTGGCTAATGTACT
GTTTGCCCCGGGAGCTCGCCAACCAGCTTGAGGCCACTGGCGTCACCTGCTATGCAGCCCACC
CAGGGCCTGTGAACTCGGAGCTGTTCTGCGCCATGTTCTTGATGGCTGCGCCCACTTTTG
CGCCCATTTGGCTTGGCTGGTGCTCCGGGACCAAGAGGGGGTGCCAGACACCCCTGTATTG
TGCTCTACAAGAGGGCATCGAGCCCCCTCAGTGGGAGATATTTTGCCAACTGCCATGTGGAAG
AGGTGCCTCCAGCTGCCCCGAGACGACCGGGCAGCCCATCGGCTATGGGAGGCCAGCAAGAGG
CTGGCAGGGCTTGGGCCTGGGGAGGATGCTGAACCCGATGAAGACCCCCAGTCTGAGGACTC
AGAGGCCCCATCTTCTCTAAGCACCCCCCACCCTGAGGAGCCCACAGTTTCTCAACCTTACC
CCAGCCCTCAGAGCTCACCAGATTTGTCTAAGATGACGCACCGAATTCAGGCTAAAGTTGAG
CCTGAGATCCAGCTCTCCTAACCCTCAGGCCAGGATGCTTGCCATGGCACCTTCATGGTCCCT
GAAAACCTCGGATGTGTGTGAGGCCATGCCCTGGACACTGACGGGTTTGTGATCTTGACCTC
CGTGGTTACTTTCTGGGGCCCCAAGCTGTGCCCTGGACATCTCTTTTCTGGTTGAAGGAAT
AATGGGTGATTATTTCTTCTGAGAGTGACAGTAACCCAGATGGAGAGATAGGGGTATGCT
AGACACTGTGCTTCTCGGAAATTTGGATGTAGTATTTTCAGGCCCCACCCTTATTGATTCTG
ATCAGCTCTGGAGCAGAGGCAGGGAGTTTGCAATGTGATGCACTGCCAACATTGAGAATTAG
TGAACCTGATCCCTTTGCAACCGTCTAGCTAGGTAGTTAAATTACCCCATGTTAATGAAGCG
GAATTAGGCTCCCGAGCTAAGGGACTCGCCTAGGGTCTCACAGTGAGTAGGAGGAGGGCCTG
GGATCTGAACCCAAGGGTCTGAGGCCAGGGCCGACTGCCGTAAGATGGGTGCTGAGAAGTGA
GTCAGGGCAGGGCAGCTGGTATCGAGGTGCCCCATGGGAGTAAGGGGACGCCTTCGGGGCGG
ATGCAGGGCTGGGGTCATCTGTATCTGAAGCCCTCGGAATAAAGCGCGTTGACCGCCAAAA
AAAAAAAAAAAAAAAAAAAA

FIGURE 75

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48227

<subunit 1 of 1, 377 aa, 1 stop

<MW: 40849, pI: 7.98, NX(S/T): 0

MEALLLGAGLLLGAYVLVYYNLVKAPPCGGMGNLRGRTAVVTGANSIGKMTALELARRGAR
VVLACRSQERGEAAAFDLRQESGNNEVIFMALDLASLASVRAFATAFLSSEPRLDILIHNA
ISSCGRTREAFNLLLRVNHIGPFLLTHLLLPCLKACAPSRVVVVASAAHCRGRLDLDFKRLDRP
VVGWRQELRAYADTKLANVLFARELANQLEATGVTCTAAHPGPVNSEFLRHVPGWLRPLLR
PLAWLVLRAPRGGAQTPLYCALQEGIEPLSGRYFANCHVEEVPPAARDRAHRLWEASKRL
AGLGPGEDAEPDEDPOSEDSEAPSSLSTPHPEEPTVSQPYPSPQSSPDLSKMTHRIQAKVEP
EIQLS

Important features:

Signal peptide:

amino acids 1-16

Glycosaminoglycan attachment site.

amino acids 46-49

Short-chain alcohol dehydrogenase family

amino acids 37-49 and 114-124

[illegible]

GAGGAGACAGCCTCCTGGGGGGGAGGGGTTCCCTGCCTCTGCTGCTCCTGCTCAT**CATGGGAGGCATGGCTCAG**
 GACTCCCCGCCCCAGATCCTAGTCCACCCCCAGGACCAGCTGTTCCAGGGCCCTGGCCCTGCCAGGATGAGCTGC
 CAAGCCTCAGGCCAGCCACCTCCCACCATCCGCTGGTTGCTGAATGGGCAGCCCCCTGAGCATGGTGCCCCCAGAC
 CCACACCCACTCCTGCCTGATGGGAGACCCCTTCTGCTGCTACAGCCCCCTGCCCGGGGACATGCCACGATGGCCAG
 GCCCTGTCCACAGCTTGGTGTCTACACATGTGAGGCCACCAACCGGCTTGGCAGCGGCTCAGCAGAGGCCGCT
 CGGCTGTCTGTGGCTGTCTCCCGGGAGGATTTCCAGATCAGCCCTCGGGACATGGTGAGCTGTGGTGGGTGAGCAG
 TTTACTCTGGAATGTGGGCCGCCCTGGGGCCACCCAGAGCCACAGTCTCATGGTGGAAAAGATGGGAAAACCCCTG
 GCCCTCCAGCCCGGAAGGCACACAGTGTCCGGGGGGTCCCTGCTGATGGCAAGAGCAGAGAAGAGTGACGAAGGG
 ACCTACATGTGTGTGGCCACCAACAGCGCAGGACATAGGGAGAGCCGCGCAGCCCGGTTTCCATCCAGGAGCC
 CAGGACTACACGGAGCCTGTGGAGCTTCTGGCTGTGCGAATTCAGCTGGAAAAATGTGACACTGCTGAACCCGGAT
 CCTGCGAGGGGCCCCAAGCCTAGACCGCGGTGTGGTTCAGTGGGAAGGTAGTGGCCCTGTGCGCTGGCCCAA
 TCTTACACGGCCTTGTTTCAGGACCCAGACTGCCCCGGGAGGCCAGGGAGCTCCGTGGGCAGAGAGCCTGCTGGCC
 GGCTGGCAGAGCGCAGAGCTTGGAGGGCTCCACTGGGGCCAAAGACTACGAGTTCAAAGTGAGACCATCCTCTGGC
 CGGGCTCGAGGCCCTGACAGCAACGTGCTGCTCCTGAGGCTGCCGGAAAAAGTGCCCAAGTGCCCCACCTCAGGAA
 GTGACTCTAAAAGCCTGGCAATGGCAGCTGTCTTTGTGAGCTGGGTCCCACCACCTGCTGAAAACCCACAAATGGCATC
 ATCCGTGGCTACAGCTTGGAGCCTGGGACCAACACATCACTGCCACCAACCTGGACTGTAGTTGGTGAGCAG
 ACCCAGCTGGAAATCGCAACCCATATGCCAGGCTCTACTGCTGCCAAGTGGCTGACATCACTGGTGCTGGAGCT
 GGGGAGCCAGTAGACCTGTCTGCCTCCTTTTATAGAGCAGGCCATGGAGCGAGCCACCCAAAGAACCCAGTGAGCAT
 GGTCCCTGGACCCTGGAGCAGCTGAGGGCTACCTTGAAGCGCCTGAGGTCAATTGCCACCTGCGGTGTTGCACTC
 TGGCTGCTGCTTCTGGGCACCGCCGTGTGTATCCACCGCCGCGCAGCTAGGGTGACCTGGGCCCAGGTCTG
 TACAGATATACCAAGTGGAGGATGCCATCCTAAAAACAGGATGGATCACAGTGACTCCCAGTGGTTGGCAGACACT
 TGGCGTTTCCACCTCTGGCTCTCGGGACCTGAGCAGCAGCAGCCTCAGCAGCTCGGTGGGGCCGGATGCCCG
 GACCCACTAGACTGTGCTGCTCCTTGTCTCTCTGGACTCCCGAAGCCCGCGCTGCCCTGCTTTCAGACACC
 AGCACTTTTTATGGCTCCCTCATCGCTGAGCTGCCCTCCAGTACCCAGCCAGGCCAAGTCCCCAGGTCCCAGCT
 GTCAGGCGCCTCCACCCAGCTGGGCCAGCTCTCCAGCCCTGTTCCAGCTCAGACAGCCTCTGCAGCCGAG
 GGACTCTCTTCTCCCCGCTTGCTCTCTGGCCCCTGACAGGCTTGGAAGGCCAAAAAGAACAGGAGCTGACAGTAT
 GCCAACAGTTCCCCACTGCTCCGGGACGCCACTCCTTGGAGCTCCGGGCTGTGAGTTAGGAAATAGAGGTTCC
 AAGAACCTTTCCCAAAGGCCAGGAGCTGTGCCCAAGCTCTGGTTGCTTGGCGGGGCTTGGGACCGAAACTCCTC
 AGCTCCTCAAATGAGCTGGTTACTCTGTCATCTCCCTCCAGCACCCCTCTTTCTCTCATGAAACTCCCCCAACTCAG
 AGTCAACAGACCCAGCCTCCGGTGGCACCACAGGCTCCCTCCTCCATCCTGCTGCCAGCAGCCCCCATCCCCATC
 CTAGACCCCTGCAGTCCCCCTAGCCCCCAGGCTCTTCCCTCTCTGGCCCCAGCCAGCTTCCAGTGCCTGTCC
 AGCTCCTCAGCTGTATCCTTGGGGAGGATCAAGACAGCGTGTGACCCCTGAGGAGGTAGCCCTGTGCTTGGAA
 CTCAGTGAGGCTGAGGAGACTCCCAAGAACAGCGTCTTCCATGCCAAGGCTCCTTACCCCCCACCCATAT
 GGGTACATCAGCGTCCCAACAGCCTCAGAGTTCAACGACATGGGCAGGACTGGAGGAGGGGTGGGGCCCAAGGGG
 GGAGTCTTGCTGTGCCACCTCGGCCCTGCCTCACCCCCACCCCCAGCGAGGGCTCCTTAGCCAATGGTTGGGG
 TCAGCCTCTGAGGACAATGCCGCCAGCGCCAGAGCCAGCCTTGTGAGTCTCTCCGATGGTCTCTTCTCGCTGAT
 GCTCACTTTTGGCCCGGCCCTGGCAGTGGCTGTGGATAGCTTTGGTTTCGGTCTAGAGCCAGGGAGGCAGACTGC
 GTCTTATAGATGCCTCATCACTTCTCCCAAGGATGAGATCTTCTGACCCCCAACCTCTCCTTGCCTTGGC
 TGGAGTGGAGGCCAGACTGGTTGGAAGACATGGAGGTGAGCCACACCCAGCGGCTGGGAAGGGGGATGCTCCCT
 TGGCCCCCTGACTCTCAGATCTCTTCCAGAGAAAGTCACTCCACTGTGCTATGCCCAAGGCTGGTGTCTTCTCCT
 GTAGATTACTCT**TGA**ACCGTGTCCCTGAGACTTCCAGACGGGAATCAGAACCCTTCTCCTGTCCACCCACAAG
 ACCTGGGCTGTGGTGTGTGGGTCTTGGCCTGTGTTTCTCTGCAGCTGGGGTCCACCTTCCCAAGCCTCCAGAGAG
 TTCTCCCTCCACGATTGTGAAAACAAATGAAAACAAATTAGAGCAAAAGCTGACCTGGAGCCCTCAGGGAGCAAA
 ACATATCTCCAGCTGACTCTTAGCAGCTGCTTCTCCTGTGCCATCCACTCCACACCAAGGTGTTTGGC
 CTGAGGAGCAGCCCTGCCCTGCTGCTCTTCCCCACCAATTTGGATCACAGGAAGTGAGGAGGCCAGAGGTGCTCTT
 GTGGAGGACAGCAGTGGCTGCTGGGAGAGGGCTGTGGAGGAAGGAGCTTCTCGAGCCCCCTCTCAGCCTTACCT
 GGGCCCCCTCCTCTAGAGAAGAGCTCAACTCTCTCCCAACCTCACCATGGAAAGAAAATAATTATGAATGCCACTG
 AGGCAGCTGAGGCCCTACTCATGCCAAAACAAAGGGTTCAAGGCTGGGTCTAGCGAGGATGCTGAAGGAAGGGAGG
 TATGACACCGTAGGTCAAAGACCACTCTCGTATGTTGTCTACTATGAGCTTAAGAAATTTGATACCATAAAAT
 GGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 77

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA41404

<subunit 1 of 1, 985 aa, 1 stop

<MW: 105336, pI: 6.55, NX(S/T): 7

MGGMAQDSPPQILVHPQDQLFQGPGPARMSCQASGQPPPTIRWLLNGQPLSMVPPDPHLLP
DGTLLLLQPPARGHAHDGQALSTD LGVYTCEASNRLGTAVSRGARLSVAVLREDFQIQPRDM
VAVVGEQFTLECGPPWGHPEPTVSWWKDGKPLALQPGRHTVSGGSLLMARAESDEGTYMCV
ATNSAGHRESRAARVSIQEPQDYTEPVELLAVRIQLENTLLNPDPAEGPKPRPAVWLSWKV
SGPAAPAQSYTALFRTQTAPGGQGAPWAEELLAGWQSAELGGLHWGQDYEFKVRPSSGRARG
PDSNVLLLLRLPEKVPSAPPQEVTLKPGNGTVFVSWVPPPAENHNGIIRGYQVWSLGNTSLPP
ANWTVVGEQTQLEIATHMPGSYCVQVAAVTGAGAGEPSRPVCLLLEQAMERATQEPSEHGWP
TLEQLRATLKRPEVIATCGVALWLLLLGTAVCIHRRRRARVHLGPGLYRTSEDAILKHRMD
HSDSQWLADTWRSTSGSRDLSSSSSLSSRLGADARDPLDCRRSLLSWDSRSPGVPLLPDTST
FYGSLIAELPSSTPARPSPQVPAVRRLPPQLAQLSSPCSSSDSLCSRRGLSSPRLSLAPAEA
WKAKKKQELQHANSPLLRGSHSLELRACELGNRGSKNLSQSPGAVPQALVAWRALGPKLLS
SSNELVTRHLPPAPLFPHETPPTQSQQQTQPPVAPQAPSSILLPAAPIPILSPCSPSPQASS
LSGPSPASSRLSSSSSLSSLGEDQDSVLTPEEVALCLELSEGEETPRNSVSPMPRAPSPPTTY
GYISVPTASEFTDMGRTGGGVGPKGGVLLCPPRPCLTPTPSEGLANGWGSASEDNAASARA
SLVSSSDGSFLADAHFARALAVAVDSFGFGLPREADCVFIDASSPPSPRDEIFLTPNLSLP
LWEWRPDWLEDMEVSHTQRLGRGMPPWPPDSQISSQRSQLHCRMPKAGASPV DYS

Important features:

Transmembrane domain:

amino acids 448-467

N-glycosylation sites:

amino acids 224-227, 338-341, 367-370, 374-377, 658-661 and 926-929

N-myristoylation sites.

amino acids 47-52, 80-85, 88-93, 99-104, 105-110, 181-186, 272-277, 290-295, 355-360, 403-408, 462-467, 561-566, 652-657, 849-854 and 876-881

Phosphotyrosine interaction domain proteins

amino acids 740-753

[illegible][illegible]

FIGURE 79

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA44196

<subunit 1 of 1, 332 aa, 1 stop

<MW: 36143, pI: 5.89, NX(S/T): 1

MRLLVLLWGCLLLPGYEALEGPTEEISGFEGDTVSLQCTYREELRDHRKYWCRKGGILFSRCS
GTIYAEEEGQETMKGRVSIRDSTRQELSLIVTLWNLTLQDAGEYWCGVEKRGPDSELLISLFV
FPGPCCPPSPSPTFQPLATTRLQPKAKAQQTQPPGLTSPGLYPAATTAKQGKTGAEAPPLPG
TSQYGHERTSQYTGTSPHPATSPAGSSRPPMQLDSTSAEDTSPALSSGSSKPRVSIPMVRI
LAPVLVLLSLLSAAGLIAFCSHLLLWRKEAQQATETQRNEKFWLSRLTAEKEAPSQAPEGD
VISMPPLHTSEEELGFSKFVSA

Important features:

Signal peptide:

amino acids 1-17

Transmembrane domain:

amino acids 248-269

N-glycosylation site.

amino acids 96-99

Fibrinogen beta and gamma chains C-terminal domain.

amino acids 104-113

Ig like V-type domain:

amino acids 13-128

FIGURE 80

TTGTGACTAAAAGCTGGCCTAGCAGGCCAGGGAGTGCAGCTGCAGGCGTGGGGGTGGCAGGA
GCCGCAGAGCCAGAGCAGACAGCCGAGAAACAGGTGGACAGTGTGAAAGAACCAGTGGTCTC
GCTCTGTTGCCCAGGCTAGAGTGTACTGGCGTGATCATAGCTCACTGCAGCCTCAGACTCCT
GGACTTGAGAAATCCTCCTGCCTTAGCCTCCTGCATATCTGGGACTCCAGGGGTGCACTCAA
GCCCTGTTTCTTCTCCTTCTGTGAGTGGACCACGGAGGCTGGTGAGCTGCCTGTCATCCCAA
AGCTCAGCTCTGAGCCAGAGTGGTGGTGGCTCCACCTCTGCCGCCGGCATAGAAGCCAGGAG
CAGGGCTCTCAGAAGGCGGTGGTGCCAGCTGGGATCATGTTGTTGGCCCTGGTCTGTCTGC
TCAGCTGCCTGCTACCCTCCAGTGAGGCCAAGCTCTACGGTCGTTGTGAACTGGCCAGAGTG
CTACATGACTTCGGGCTGGACGGATACCGGGGATACAGCCTGGCTGACTGGGTCTGCCTTGC
TTATTTTACAAGCGGTTTCAACGCAGCTGCTTTGGACTACGAGGCTGATGGGAGCACCAACA
ACGGGATCTTCCAGATCAACAGCCGGAGGTGGTGCAGCAACCTCACCCCGAACGTCCCCAAC
GTGTGCCGGATGTACTGCTCAGATTTGTTGAATCCTAATCTCAAGGATACCGTTATCTGTGC
CATGAAGATAACCCAAGAGCCTCAGGGTCTGGGTACTGGGAGGCCTGGAGGCATCACTGCC
AGGGAAAAGACCTCACTGAATGGGTGGATGGCTGTGACTTCTTAGGATGGACGGAACCATGCA
CAGCAGGCTGGGAAATGTGGTTTGGTTCCTGACCTAGGCTTGGGAAGACAAGCCAGCGAATA
AAGGATGGTTGAACGTGAAA

bioRxiv preprint doi: <https://doi.org/10.1101/2020.03.10.336000>; this version posted March 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 81

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA52187

<subunit 1 of 1, 146 aa, 1 stop

<MW: 16430, pI: 5.05, NX(S/T): 1

MLLALVCLLSCLLPSSSEAKLYGRCELARVLHDFGLDGYRGYSLADWVCLAYFTSGFNAAALD
YEADGSTNNGIFQINSRRWCSNLTPNVPNVCRMYSDDLNPNLKDTVICAMKITQEPQGLGY
WEAWRHHCQGKDLTEWVDGPDF

Important features:

Signal peptide:

amino acids 1-18

N-myristoylation site.

amino acids 67-72

Homologous region to Alpha-lactalbumin / lysozyme C proteins.

amino acids 34-58 (catalytic domain), 111-132 and 66-107

[illegible][illegible]

FIGURE 83

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48328

<subunit 1 of 1, 351 aa, 1 stop

<MW: 39052, pI: 8.97, NX(S/T): 2

MSPRSCLRSLRLLVFAVFSAAASNWLYLAKLSSVGSISEEETCEKLKGLIQRQVQMCKRNLE
VMDSVRRGAQLAIEECQYQFRNRRWNCSTLDSL PVFGKVVTQGTREAAFVYA ISSAGVAFV
TRACSSGELEKCGCDRTVHGVSPQGFQWSGCSDNIAYGVAFS QSFVDVRERSKGASSSRALM
NLHNNEAGRKAILTHMRVECKCHGVSGSCEVKTCWRAVPPFRQVGHALKEKFDGATEVEPRR
VGSSRALVPRNAQFKPHTDEDLVYLEPSPDFCEQDMRSGVLGTRGRTC NKTSKAIDGCELLC
CGRGFHTAQVELAERCSCKFHWCCFVKCRQCQRLVELHTCR

Important features:

Signal peptide:

amino acids 1-22

N-glycosylation sites.

amino acids 88-91 and 297-300

Wnt-1 family signature.

amino acids 206-215

Homologous region to Wnt-1 family proteins

amino acids 183-235, 305-350, 97-138, 53-92 and 150 -174

FIGURE 84

CGGACGCGTGGGCGGACGCGTGGGCGGACGCGTGGGCGGACGCGTGGGCTGGGTGCCTGCAT
CGCCATGGACACCACCAGGTACAGCAAGTGGGGCGGCAGCTCCGAGGAGGTCCCCGGAGGGC
CCTGGGGACGCTGGGTGCACTGGAGCAGGAGACCCCTCTTCTTGGCCCTGGCTGTCCTGGTC
ACCACAGTCCTTTGGGCTGTGATTCTGAGTATCCTATTGTCCAAGGCCTCCACGGAGCGCGC
GGCGCTGCTTGACGGCCACGACCTGCTGAGGACAAACGCCTCGAAGCAGACGGCGGCGCTGG
GTGCCCTGAAGGAGGAGGTCTGGAGACTGCCACAGCTGCTGCTCGGGGACGCAGGCGCAGCTG
CAGACCACGCGCGCGGAGCTTGGGGAGGCGCAGGCGAAGCTGATGGAGCAGGAGAGCGCCCT
GCGGGAACTGCGTGAGCGCGTGACCCAGGGCTTGGCTGAAGCCGCGCAGGGGGCCGTGAGGACG
TCCGCACTGAGCTGTTCCGGGCGCTGGAGGGCCGTGAGGCTCCAGAACAACCTCCTGCGAGCCG
TGCCCCACGTGCTGGCTGTCTTCGAGGGCTCCTGCTACTTTTTCTCTGTGCCAAAGACGAC
GTGGGCGGCGGCGCAGGATCACTGCGCAGATGCCAGCGCGCACCTGGTGATCGTTGGGGGGCC
TGGATGAGCAGGGCTTCCTCACTCGGAACACGCGTGGCCGTGGTTACTGGCTGGGCGCTGAGG
GCTGTGCGCCATCTGGGCAAGGTTTCAGGGCTACCACTGGGTGGACGGAGTCTCTCTCAGCTT
CAGCCACTGGAACCAGGGAGAGCCCAATGACGCTTGGGGGCGCGAGAAGTGTGTCATGATGC
TGCACACGGGGCTGTGGAACGACGCACCGTGTGACAGCGAGAAGGACGGCTGGATCTGTGAG
AAAAGGCACAACCTGCTTGACCCCGCCAGTGCCCTGGAGCCGCGCCCATTCAGCATGTCGTA
TCCTGGGGGCTGCTCACCTCCCTGGCTCCTGGAGCTGATTGCCAAAGAGTTTTTTTCTTCCT
CATCCACCGCTGCTGAGTCTCAGAAACACTTGGCCCAACATAGCCCTGTCCAGCCCAGTGCC
TGGGCTCTGGGACCTCCATGCCGACCTCATCCTAACTCCACTCACGCAGACCCAACTAACC
TCCACTAGCTCCAAAATCCCTGCTCCTGCGTCCCCGTGATATGCCTCCACTTCTCTCCCTAA
CCAAGGTTAGGTGACTGAGGACTGGAGCTGTTTGGTTTTCTCGCATTTTCCACCAAACCTGGA
AGCTGTTTTTGCAGCCTGAGGAAGCATCAATAAATATTTGAGAAATGAAAAA

FIGURE 85

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56352

<subunit 1 of 1, 293 aa, 1 stop

<MW: 32562, pI: 6.53, NX(S/T): 2

MDTTRYSKWGSSEEVPGGPWGRWVHWSRRPLFLALAVLVTTVLWAVILSILLSKASTERAA
LLDGHDLRLRTNASKQTAALGALKEEVGDCHSCCSGTQAQLQTTRAELGEAQAKLMEQESALR
ELRERVLTQGLAEAGRGREDVRTELFRALAVRLQNNSCPCPTSWLSFEGSCYFFSVPKTTW
AAAQDHCADASAHLVIVGGLDEQGFLTRNTRGRGYWLGLRAVRHLGKVQGYQWVDGVSLSFS
HWNQGEPNDAWGRENCVMMMLHTGLWNDAPCDSEKDGWICEKRHNC

Important features:

Type II transmembrane domain:

amino acids 31-54

N-glycosylation sites.

amino acids 73-76 and 159-162

Leucine zipper pattern.

amino acids 102-123

N-myristoylation sites.

amino acids 18-23, 133-138 and 242-247

C-type lectin domain signature.

amino acids 264-287

FIGURE 86

GCCAGGGGAAGAGGGTGATCCGACCCGGGGAAGGTCGCTGGGCAGGGCGAGTTGGGAAAGCG
GCAGCCCCCGCCGCCCCCGCAGCCCCTTCTCCTCCTTTCTCCACGTCCTATCTGCCTCTCG
CTGGAGGCCAGGCCGTGCAGCATCGAAGACAGGAGGAACTGGAGCCTCATTGGCCGGCCCCG
GGCGCCGGCCTCGGGCTTAAATAGGAGCTCCGGGCTCTGGCTGGGACCCGACCGCTGCCGGC
CGCGCTCCCGCTGCTCCTGCCGGGTGATGGAAAACCCAGCCCCGGCCCGCCCTGGGCAAG
GCCCTCTGCGCTCTCCTCCTGGCCACTCTCGGCGCCGCGGCCAGCCTCTTGGGGGAGAGTC
CATCTGTTCCGCCAGAGCCCCGGCCAAATACAGCATCACCTTCACGGGCAAGTGAGCCAGA
CGGCCTTCCCCAAGCAGTACCCCTGTTCCGCCCCCTGCGCAGTGGTCTTCGCTGCTGGGG
GCCGCGCATAGCTCCGACTACAGCATGTGGAGGAAGAACCAGTACGTACGTAACGGGCTGCG
CGACTTTGCGGAGCGCGGCGAGGCCTGGGCGCTGATGAAGGAGATCGAGGCGGCGGGGAGG
CGCTGCAGAGCGTGACAGAGGTGTTTTTCGGCGCCCGCCGTCCCCAGCGGCACCGGGCAGACG
TCGGCGGAGCTGGAGGTGCAGCGCAGGCACTCGCTGGTCTCGTTTGTGGTGCGCATCGTGCC
CAGCCCCGACTGGTTTCGTGGGCGTGACAGCCTGGACCTGTGCGACGGGGACCGTTGGCGGG
AACAGGCGGCGCTGGACCTGTACCCCTACGACGCCGGGACGGACAGCGGCTTCACCTTCTCC
TCCCCAACTTCGCCACCATCCCGCAGGACACGGTGACCGAGATAACGTCTCTCTCCCAG
CCACCCGGCCAACTCCTTCTACTACCCGCGGCTGAAGGCCCTGCCTCCCATCGCCAGGGTGA
CACTGCTGCGGCTGCGACAGAGCCCCAGGGCCTTCATCCCTCCCGCCCCAGTCCTGCCCAGC
AGGGACAATGAGATTGTAGACAGCGCCTCAGTTCCAGAAACGCCGCTGGACTGCGAGGTCTC
CCTGTGGTCGTCTGGGGACTGTGCGGAGGCCACTGTGGGAGGCTCGGGACCAAGAGCAGGA
CTCGCTACGTCCGGGTCCAGCCCCGCAACAACGGGAGCCCCCTGCCCCGAGCTCGAAGAAGAG
GCTGAGTGCGTCCCTGATAACTGCGTCTAAGACCAGAGCCCCCGCAGCCCCCTGGGGCCCCCG
GAGCCATGGGGTGTGCGGGGCTCCTGTGCAGGCTCATGCTGCAGGCGGCCGAGGGCACAGGG
GGTTTCGCGCTGCTCCTGACCGCGGTGAGGCCGCGCCGACCATCTCTGCACTGAAGGGCCCT
CTGGTGGCCGGCACGGGCATTGGGAAACAGCCTCCTCCTTTCCCAACCTTGCTTCTTAGGGG
CCCCCGTGTCCCGTCTGCTCTCAGCCTCCTCCTCCTGCAGGATAAAGTCATCCCCAAGGCTC
CAGCTACTCTAAATTATGTCTCCTTATAAGTTATTGCTGCTCCAGGAGATTGTCTTTCATCG
TCCAGGGGCTGGCTCCCACGTGGTTGCAGATACCTCAGACCTGGTGCTCTAGGCTGTGCTG
AGCCCCTCTCCCGAGGGCGCATCCAAGCGGGGGCCACTTGAGAAGTGAATAAATGGGGCGG
TTTCGGAAGCGTCAGTGTTTCCATGTTATGGATCTCTCTGCGTTTGAATAAAGACTATCTCT
GTTGCTCACAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

Variable	Mean	SD	Min	Max
Age	38.5	12.5	25	65
Gender	0.5	0.5	0	1
Marital status	0.5	0.5	0	1
Education	12.5	2.5	9	16
Income	1500	500	500	3000
Health status	0.5	0.5	0	1
Smoking status	0.5	0.5	0	1
Alcohol consumption	0.5	0.5	0	1
Exercise frequency	0.5	0.5	0	1
Stress level	0.5	0.5	0	1
Sleep quality	0.5	0.5	0	1
Work satisfaction	0.5	0.5	0	1
Life satisfaction	0.5	0.5	0	1
Overall health	0.5	0.5	0	1
Depression score	0.5	0.5	0	1
Anxiety score	0.5	0.5	0	1
Stress score	0.5	0.5	0	1
Sleep score	0.5	0.5	0	1
Work score	0.5	0.5	0	1
Life score	0.5	0.5	0	1
Health score	0.5	0.5	0	1
Depression score	0.5	0.5	0	1
Anxiety score	0.5	0.5	0	1
Stress score	0.5	0.5	0	1
Sleep score	0.5	0.5	0	1
Work score	0.5	0.5	0	1
Life score	0.5	0.5	0	1
Health score	0.5	0.5	0	1
Depression score	0.5	0.5	0	1
Anxiety score	0.5	0.5	0	1
Stress score	0.5	0.5	0	1
Sleep score	0.5	0.5	0	1
Work score	0.5	0.5	0	1
Life score	0.5	0.5	0	1
Health score	0.5	0.5	0	1
Depression score	0.5	0.5	0	1
Anxiety score	0.5	0.5	0	1
Stress score	0.5	0.5	0	1
Sleep score	0.5	0.5	0	1
Work score	0.5	0.5	0	1
Life score	0.5	0.5	0	1
Health score	0.5	0.5	0	1
Depression score	0.5	0.5	0	1
Anxiety score	0.5	0.5	0	1
Stress score	0.5	0.5	0	1
Sleep score	0.5	0.5	0	1
Work score	0.5	0.5	0	1
Life score	0.5	0.5	0	1
Health score	0.5	0.5	0	1
Depression score	0.5	0.5	0	1
Anxiety score	0.5	0.5	0	1
Stress score	0.5	0.5	0	1
Sleep score	0.5	0.5	0	1
Work score	0.5	0.5	0	1
Life score	0.5	0.5	0	1
Health score	0.5	0.5	0	1
Depression score	0.5	0.5	0	1
Anxiety score	0.5	0.5	0	1
Stress score	0.5	0.5	0	1
Sleep score	0.5	0.5	0	1
Work score	0.5	0.5	0	1
Life score	0.5	0.5	0	1
Health score	0.5	0.5	0	1
Depression score	0.5	0.5	0	1
Anxiety score	0.5	0.5	0	1
Stress score	0.5	0.5	0	1
Sleep score	0.5	0.5	0	1
Work score	0.5	0.5	0	1
Life score	0.5	0.5	0	1
Health score	0.5	0.5	0	1
Depression score	0.5	0.5	0	1
Anxiety score	0.5	0.5	0	1
Stress score	0.5	0.5	0	1
Sleep score	0.5	0.5	0	1
Work score	0.5	0.5	0	1
Life score	0.5	0.5	0	1
Health score	0.5	0.5	0	1
Depression score	0.5	0.5	0	1
Anxiety score	0.5	0.5	0	1
Stress score	0.5	0.5	0	1
Sleep score	0.5	0.5	0	1
Work score	0.5	0.5	0	1
Life score	0.5	0.5	0	1
Health score	0.5	0.5	0	1
Depression score	0.5	0.5	0	1
Anxiety score				

```
><subunit 1 of 1, 331 aa, 1 stop
```

MENPSPAAALGKALCALLLATLGAAGQPLGGESIC SARAPAKYSITFTGKWSQTAFPKQYPL
FRPPAQWSSLLGAHSSDYSMWRKNQYVSNGLRDF AERGEAWALMKEIEAAGEALQSVHEVF
SAPAVPSGTGQTSAELEVQRRHSLVSFVVRIVPSPDWFGVDSLDCGDRWREQAALDLYP
YDAGTDSGFTFSSPNFATIPQDTVTEITSSSPSHPAN SFYYPRLKALPPIARVTLLRLRQSP
RAFIPPAVLPSRDNEIVDSASVPETPLDCEVSLWSSWGLCGGHCGR LGTKSRTRYVRVQPA
NNGSPCPELEEEAECPDNCV

Signal peptide:

amino acids 1-26

FIGURE 88

GGCGGCGTCCGTGAGGGGCTCCTTTGGGCAGGGGTAGTGTGGTGTCCCTGTCTTGCGTGA
TATTGACAAACTGAAGCTTTCCTGCACCACTGGACTTAAGGAAGAGTGTACTCGTAGGCGGA
CAGCTTTAGTGGCCGGCCGGCCGCTCTCATCCCCGTAAGGAGCAGAGTCCTTTGTACTGAC
CAAGATGAGCAACATCTACATCCAGGAGCCTCCCACGAATGGGAAGGTTTTATTGAAAACCTA
CAGCTGGAGATATTGACATAGAGTTGTGGTCCAAAGAAGCTCCTAAAGCTTGCAGAAATTTT
ATCCAACCTTTGTTTTGGAAGCTTATTATGACAATACCATTTTTTCATAGAGTTGTGCCTGGTTT
CATAGTCCAAGGCGGAGATCCTACTGGCACAGGGAGTGGTGGAGAGTCTATCTATGGAGCGC
CATTCAAAGATGAATTTCAATCACGGTTGCGTTTTAATCGGAGAGGACTGGTTGCCATGGCA
AATGCTGGTTCTCATGATAATGGCAGCCAGTTTTTCTTCACACTGGGTGAGCAGATGAACT
TAACAATAAGCATAACCATCTTTGGAAAGGTTACAGGGGATACAGTATATAACATGTTGCGAC
TGTGAGAAGTAGACATTGATGATGACGAAAGACCACATAATCCACACAAAATAAAAAGCTGT
GAGGTTTTGTTTAATCCTTTTGATGACATCATTCCAAGGGAAATTAAAAGGCTGAAAAAGA
GAAACCAGAGGAGGAAGTAAAGAAATTGAAACCCAAAGGCACAAAAAATTTAGTTTACTTT
CATTTGGAGAGGAAGCTGAGGAAGAAGAGGAGGAAGTAAATCGAGTTAGTCAGAGCATGAAG
GGCAAAGCAAAGTAGTCATGACTTGCTTAAGGATGATCCACATCTCAGTTCTGTTCCAGT
TGTAGAAAGTGAAAAAGGTGATGCACCAGATTTAGTTGATGATGGAGAAGATGAAAGTGCAG
AGCATGATGAATATATTGATGGTGATGAAAAGAACCTGATGAGAGAAAGAATTGCCAAAAA
TTAAAAAAGGACACAAGTGCGAATGTTAAATCAGCTGGAGAAGGAGAAGTGGAGAAGAAATC
AGTCAGCCGCAGTGAAGAGCTCAGAAAAGAAGCAAGACAATTAAACGGGAACCTTTAGCAG
CAAAACAAAAAAGTAGAAAATGCAGCAAAACAAGCAGAAAAAGAAGTGAAGAGGAAGAA
GCCCCCTCCAGATGGTGCTGTTGCCGAATACAGAAGAGAAAAAGCAAAGTATGAAGCTTTGAG
GAAGCAACAGTCAAAGAAGGGAACCTCCCGGAAGATCAGACCCTTGCACTGCTGAACCAGT
TTAAATCTAACTCACTCAAGCAATTGCTGAAACACCTGAAAATGACATTCCTGAAACAGAA
GTAGAAGATGATGAAGGATGGATGTCACATGTACTTCAGTTTGAGGATAAAAGCAGAAAAGT
GAAAGATGCAAGCATGCAAGACTCAGATACATTTGAAATCTATGATCCTCGGAATCCAGTGA
ATAAAAGAAGGAGGGAAGAAAGCAAAAAGCTGATGAGAGAGAAAAAAGAAAGAAGATATAAAT
GAGAATAATGATAACCAGAACTTGCTGGAAATGTGCCTACAATGGCCTTGTAACAGCCATTG
TTCCAACAGCATCACTTAGGGGTGTGAAAAGAAGTATTTTTGAACCTGTTGTCTGGTTTTG
AAAAACAATTATCTTGTGTTTTGCAAATTGTGGAATGATGTAAGCAAATGCTTTTGGTTACTGG
TACATGTGTTTTTTCCTAGCTGACCTTTTATATTGCTAAATCTGAAATAAAATAACTTTCCT
TCCACAAAAA

[illegible]

><MW: 53847, pI: 5.75, NX(S/T): 2

amino acids 96-140, 49-89 and 22-51

FIGURE 90

CGCCGCCGTTGGGGCTGGAAGTTCCCGCCAGGTCCGTGCCGGGCGAGAGAGATGCTGCCCGG
CCCGCCTCGGCTTTGAGGCGAGAGAAGTGTCCCAGACCCATTTTCGCCTTGCTGACGGCGTGC
AGCCCTGGCCAGACATGTCCACAGGGTTCTCCTTCGGGTCCGGGACTCTGGGCTCCACCACC
GTGGCCGCGGCGGGACCAGCACAGGCGGCGTTTTCTCCTTCGGAACGGGAACGTCTAGCAA
CCCTTCTGTGGGGCTCAATTTTGGAATCTTGGAAGTACTTCAACTCCAGCAACTACATCTG
CTCCTTCAAGTGTTTTTGGAACCGGGCTCTTTGGATCTAAACCTGCCACTGGGTTCACTCTA
GGAGGAACAAATACAGGTGCCCTTGACACCAAGAGGCCTCAAGTGGTCACCAAATATGGAAC
CCTGCAAGGAAAACAGATGCATGTGGGGAAGACACCCATCCAAGTCTTTTTAGGAGTCCCCT
TCTCCAGACCTCCTCTAGGTATCCTCAGGTTTGCACCTCCAGAACCCCCGGAGCCCTGGAAA
GGAATCAGAGATGCTACCACCTACCCGCCTGGATGGAGTCTCGCTCTGTCGCCAGGCTGGAG
TGCAGTGGCACGATCTCGGCTCACTGCAACCTCCGCCTCCCGGGTTCAAGCGAGTCTCCTGC
CTCAGCCTCTGAGTGTCTGGGGCTACAGGTGCCCTGCAGGAGTCTGGGGCCAGCTGGCCTCG
ATGTACGTACGACGCGGGAACGGTACAAGTGGCTGCGCTTCAGCGAGGACTGTCTGTACCT
GAACGTGTACGCGCCGCGCGCGCGCCCGGGGATCCCCAGCTGCCAGTGATGGTCTGGTTCC
CGGGAGGCGCCTTCATCGTGGGCGCTGCTTCTTCGTACGAGGGCTCTGACTTGGCCGCCCGC
GAGAAAGTGGTGCTGGTGTTTTCTGCAGCACAGGCTCGGCATCTTCGGCTTCTTGAGCACGGA
CGACAGCCACGCGCGCGGGAACCTGGGGGCTGCTGGACCAGATGGCGGCTCTGCGCTGGGTGC
AGGAGAACATCGCAGCCTTCGGGGGAGACCCAGGAAATGTGACCCTGTTTCGGCCAGTCGGCG
GGGGCCATGAGCATCTCAGGACTGATGATGTCAACCCTAGCCTCGGGTCTCTTCCATCGGGC
CATTTCCCAGAGTGGCACCGCGTTATTACAGACTTTTCATCACTAGTAACCCACTGAAAGTGG
CCAAGAAGGTTGCCACCTGGCTGGATGCAACCACAACAGCACACAGATCCTGGTAAACTGC
CTGAGGGCACTATCAGGGACCAAGGTGATGCGTGTGTCCAACAAGATGAGATTCTTCCAAC
GAACTTCCAGAGAGACCCGGAAGAGATTATCTGGTCCATGAGCCCTGTGGTGGATGGTGTGG
TGATCCCAGATGACCCTTTGGTGCTCCTGACCCAGGGGAAGGTTTCATCTGTGCCCTACCTT
CTAGGTGTCAACAACCTGGAATTCAATTGGCTCTTGCCCTTATAATATCACCAAGGAGCAGGT
ACCACTTGTGGTGGAGGAGTACCTGGACAATGTCAATGAGCATGACTGGAAGATGCTACGAA
ACCGTATGATGGACATAGTTCAAGATGCCACTTTCGTGTATGCCACACTGCAGACTGCTCAC
TACCACCGAGAAACCCCAATGATGGGAATCTGCCCTGCTGGCCACGCTACAACAAGGATGAA
AAGTACCTGCAGCTGGATTTTACCACAAGAGTGGGCATGAAGCTCAAGGAGAAGAAGATGGC
TTTTTGGATGAGTCTGTACCAGTCTCAAAGACCTGAGAAGCAGAGGCAATTCTAAGGGTGGC
TATGCAGGAAGGAGCCAAAGAGGGGTTTGGCCCCACCATCCAGGCCCTGGGGAGACTAGCCA
TGGACATACCTGGGGACAAGAGTTCTACCCACCCAGTTTGAAGTGCAGGAGCTCCCTGCT
GCCTCCAGGCCAAAGCTAGAGCTTTTGCCTGTTGTGTGGGACCTGCACTGCCCTTTCCAGCC
TGACATCCCATGATGCCCTCTACTTCACTGTTGACATCCAGTTAGGCCAGGCCCTGTCAAC
ACCACACTGTGCTCAGCTCTCCAGCCTCAGGACAACCTCTTTTTTTTCCCTTCTTCAAATCCT
CCCACCTTCAATGTCTCCTTGTGACTCCTTCTTATGGGAGGTCGACCCAGACTGCCACTGC
CCCTGTCACTGCACCCAGCTTGGCATTACCATCCATCCTGCTCAACCTTGTTCCTGTCTGT
TCACATTGGCCTGGAGGCCTAGGGCAGGTTGTGACATGGAGCAAACCTTTTGGTAGTTTGGGA
TCTTCTCTCCACCCACACTTATCTCCCCAGGGCCACTCCAAAGTCTATACACAGGGGTGG
TCTCTTCAATAAAGAAGTGTTGATTAGAAAAA

FIGURE 91

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA44179

<subunit 1 of 1, 545 aa, 1 stop

<MW: 58934, pI: 9.45, NX(S/T): 4

MSTGFSFGSGTLGSTTVAAGGTSTGGVFSFGTGTSSNPSVGLNFGNLGSTSTPATT SAPSSG
FGTGLFGSKPATGFTLGGTNTGALHTKRPQVVTKYGT LQGKQMHVGKTPIQVFLGVPF SRPP
LGILRFAPPEPPEPWKGIRDATTYPGWSLALSPGWSAVARSRLTATSASRVQASLLPQPLS
VWGYRCLQESWGQLASMYVSTRERYKWLRFSEDCLYLNVYAPARAPGDPQLPVMVWFPGGAF
IVGAASSYEGSDLAAREKVVLVFLQHRLGIFGFLSTDDSHARGNWGLLDQMAALRWVQENIA
AFGGDPGNVT LFGQSAGAMSISGLMMSPLASGLFHRAISQSGTALFRLFITSNPLKVAKKVA
HLAGCNHNSTQILVNCLRALSGTKVMRVS NKMRFLQLNFQRDPEEIIWSMSPVVDGVVIPDD
PLVLLTQGKVSSVPYLLGVNNLEFNWLLPYNITKEQVPLVVEEYLDNVNEHDWKMLRNRMD
IVQDATFVYATLQTAHYHRET PMMGICPAGHATTRMKSTCSWILPQEWA

Important features:

Signal peptide:

amino acids 1-29

Carboxylesterases type-B serine active site.

amino acids 312-327

Carboxylesterases type-B signature 2.

amino acids 218-228

N-glycosylation sites.

amino acids 318-321, 380-383 and 465-468

FIGURE 92

GAGAACAGGCCTGTCTCAGGCAGGCCCTGCGCCTCCTATGCGGAGATGCTACTGCCACTGCT
GCTGTCCTCGCTGCTGGGCGGGTCCCAGGCTATGGATGGGAGATTCTGGATACGAGTGCAGG
AGTCAGTGATGGTGCCGGAGGGCCTGTGCATCTCTGTGCCCTGCTCTTTCTCCTACCCCCGA
CAAGACTGGACAGGGTCTACCCCAGCTTATGGCTACTGGTTCAAAGCAGTGACTGAGACAAC
CAAGGGTGCTCCTGTGGCCACAAACCACCAGAGTCGAGAGGTGGAAATGAGCACCCGGGGCC
GATTCCAGCTCACTGGGGATCCCGCCAAGGGGAACTGCTCCTTGGTGATCAGAGACGCGCAG
ATGCAGGATGAGTCACAGTACTTCTTTTCGGGTGGAGAGAGGAAGCTATGTGACATATAATTT
CATGAACGATGGGTTCTTTCTAAAAGTAACAGTGCTCAGCTTCACGCCCAGACCCCAGGACC
ACAACACCGACCTCACCTGCCATGTGGACTTCTCCAGAAAGGGTGTGAGCGCACAGAGGACC
GTCCGACTCCGTGTGGCCTATGCCCCCAGAGACCTTGTTATCAGCATTTTACAGTGACAACAC
GCCAGCCCTGGAGCCCCAGCCCCAGGGAAATGTCCCATACCTGGAAGCCCCAAAAGGCCAGT
TCCTGCGGCTCCTCTGTGCTGCTGACAGCCAGCCCCCTGCCACACTGAGCTGGGTCTGCGAG
AACAGAGTCCTCTCCTCGTCCCATCCCTGGGGCCCTAGACCCCTGGGGCTGGAGCTGCCCCG
GGTGAAGGCTGGGGATTTCAGGGCGCTACACCTGCCGAGCGGAGAACAGGCTTGGCTCCAGC
AGCGAGCCCTGGACCTCTCTGTGCAGTATCCTCCAGAGAACCTGAGAGTGATGGTTTCCCAA
GCAAACAGGACAGTCCTGGAAAACCTTGGGAACGGCACGTCTCTCCAGTACTGGAGGGCCA
AAGCCTGTGCCTGGTCTGTGTACACACAGCAGCCCCCAGCCAGGCTGAGCTGGACCCAGA
GGGGACAGGTTCTGAGCCCCCTCCAGCCCTCAGACCCCCGGGGTCTGGAGCTGCCTCGGGTT
CAAGTGGAGCACGAAGGAGAGTTACCTGCCACGCTCGGCACCCACTGGGCTCCAGCACGT
CTCTCTCAGCCTCTCCGTGCACTATAAGAAGGGACTCATCTCAACGGCATTCTCCAACGGAG
CGTTTCTGGGAATCGGCATCACGGCTCTTCTTTTCTCTGCTGCGCCTGATCATCATGAAG
ATTCTACCGAAGAGACGGACTCAGACAGAAACCCCGAGGCCAGGTTCTCCCGGCACAGCAC
GATCCTGGATTACATCAATGTGGTCCCGACGGCTGGCCCCCTGGCTCAGAAGCGGAATCAGA
AAGCCACACCAAACAGTCCTCGGACCCCTCCTCCACCAGGTGCTCCCTCCCCAGAATCAAAG
AAGAACCAGAAAAAGCAGTATCAGTTGCCAGTTTCCAGAACCCAAATCATCCACTCAAGC
CCCAGAATCCCAGGAGAGCCAAGAGGAGCTCCATTATGCCACGCTCAACTTCCCAGGCGTCA
GACCCAGGCCTGAGGCCCGGATGCCCAAGGGCACCCAGGCGGATTATGCAGAAGTCAAGTTC
CAATGAGGGGTCTCTTAGGCTTTAGGACTGGGACTTCCGGCTAGGGAGGAAGGTAGAGTAAGAG
GTTGAAGATAACAGAGTGCAAAGTTTCTTCTCTCCCTCTCTCTCTCTCTCTCTCTCTCTCT
CTCTCTTTCTCTCTCTTTTAAAAAAACATCTGGCCAGGGCACAGTGGCTCACGCCTGTAATC
CCAGCACTTTGGGAGGTTGAGGTGGGCAGATCGCCTGAGGTGGGAGTTCGAGACCAGCCTG
GCCAACTTGGTGAAACCCCGTCTCTACTAAAAATACAAAAATTAGCTGGGCATGGTGGCAGG
CGCCTGTAATCCTACCTACTTGGGAAGCTGAGGCAGGAGAATCACTTGAACCTGGGAGACGG
AGGTTGCAGTGAGCCAAGATCACACCATTGCACGCCAGCCTGGGCAACAAAGCGAGACTCCA
TCTCAAAAAAAAATCCTCCAAATGGGTTGGGTGTCTGTAATCCCAGCACTTTGGGAGGCTA
AGGTGGGTGGATTGCTTGAGCCCAGGAGTTTCGAGACCAGCCTGGGCAACATGGTGAAACCCC
ATCTCTACAAAAAATACAAAACATAGCTGGGCTTGGTGGTGTGTGCCTGTAGTCCAGCTGT
CAGACATTTAAACCAGAGCAACTCCATCTGGAATAGGAGCTGAATAAAATGAGGCTGAGACC
TACTGGGCTGCATTCTCAGACAGTGGAGGCATTCTAAGTCACAGGATGAGACAGGAGGTCCG
TACAAGATACAGGTCATAAAGACTTTGCTGATAAAACAGATTGCAGTAAAGAAGCCAACCAA
ATCCCACCAAAACCAAGTTGGCCACGAGAGTGACCTCTGGTCGTCTCACTGCTACACTCCT
GACAGCACCATGACAGTTTACAAATGCCATGGCAACATCAGGAAGTTACCCGATATGTCCCA
AAAGGGGGAGGAATGAATAATCCACCCCTTGTTTAGCAAATAAGCAAGAAATAACCATAAAA
GTGGGCAACCAGCAGCTCTAGGCGTGTCTTTGTCTATGGAGTAGCCATTCTTTTGTTCCTT
TACTTTCTTAATAAACTTGCTTTTACCTTAAAAAAA

FIGURE 93

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA54002

><subunit 1 of 1, 544 aa, 1 stop

><MW: 60268, pI: 9.53, NX(S/T): 3

MLLPLLLSSLLGGSQAMDGRFWIRVQESVMVPEGLCISVPCSFSPRQDWTGSTPAYGYWFK
AVTETTTKGAPVATNHQSREVESTRGRFQLTGDPKGNCSLVIRDAQMQDESYFFRVERGS
YVTYNFMNDGFFLKVTVLSTPRPQDHNTDLTCHVDFSRKGVSAQRTVRLRVAYAPRDLVIS
ISRDNTPALEPQPQGNVPYLEAQKGQFLRLLCAADSQPPATLSWVLQNRVLSSSHPWGPRPL
GLELPGVKAGDSGRYTCRAENRLGSQQRALDLSVQYPPENLRVMVSQANRTVLENLGNGTSL
PVLEGQSLCLVCVTHSSPPARLSWTQRGQVLSPSQSPDPGVLELPRVQVEHEGEFTCHARHP
LGSQHVSLSLSVHYKKGLISTAFSNGAFLGIGITALLFLCLALIIMKILPKRRTQTETPRPR
FSRHSTILDYINVVPTAGPLAQKRNQKATPNSPRTPPPPGAPSPESKKNQKKQYQLPSFPEP
KSSTQAPESQESQEELHYATLNFPGVRPRPEARMPKGTQADYAEVKFQ

Important features:

Signal peptide:

amino acids 1-15

Transmembrane domain:

amino acids 399-418

N-glycosylation site.

amino acids 100-103, 297-300 and 306-309

Immunoglobulins and major histocompatibility complex proteins signature.

amino acids 365-371

[illegible]

TGAGAGTAATAGTTGGAATCAAAAGAGTCAACGCAATGAACCTGTTATTTACTGCTGCGTTT
TATGTTGGGAATTCCTCTCCTATGGCCTTGTCTTGAGAGCAACAGAAAACCTCTCAAACAAAGA
AAGTCAAGCAGCCAGTGCGATCTCATTTGAGAGTGAAGCGTGGCTGGGTGTGGAACCAATTT
TTTGTACCAGAGGAAATGAATACGACTAGTCATCACATCGGCCAGCTAAGATCTGATTTAGA
CAATGGAAACAATTCCTTCCAGTACAAGCTTTTGGGAGCTGGAGCTGGAAGTACTTTTATCA
TTGATGAAAGAACAGGTGACATATATGCCATACAGAAGCTTGATAGAGAGGAGCGATCCCTC
TACATCTTAAGAGCCCAGGTAATAGACATCGCTACTGGAAGGGCTGTGGAACCTGAGTCTGA
GTTTGTCTATCAAAGTTTTCGGATATCAATGACAATGAACCAAAATTCCTAGATGAACCTTATG
AGGCCATTGTACCAGAGATGTCTCCAGAAGGAACATTAGTTATCCAGGTGACAGCAAGTGAT
GCTGACGATCCCTCAAGTGGAATAATGCTCGTCTCCTCTACAGCTTACTTCAAGGCCAGCC
ATATTTTTCTGTTGAACCAACAACAGGAGTCATAAGAATATCTTCTAAAATGGATAGAGAAC
TGCAAGATGAGTATTGGGTAATCATTTCAAGCCAAGGACATGATTGGTCAGCCAGGAGCGTTG
TCTGGAACAACAAGTGTATTAATTAACCTTTAGATGTTAATGACAATAAGCCTATATTTAA
AGAAAGTTTATACCGCTTGACTGTCTCTGAATCTGCACCCACTGGGACTTCTATAGGAACAA
TCATGGCATATGATAATGACATAGGAGAGAAATGCAGAAATGGATTACAGCATTGAAGAGGAT
GATTCGCAAACATTTGACATTATTACTAATCATGAACTCAAGAAGGAATAGTTATATTTAAA
AAAGAAAGTGGATTTTGGAGCACCAGAACCACTACGGTATTAGAGCAAAAGTTAAAAACCATC
ATGTTCTTGAGCAGCTCATGAAGTACCACACTGAGGCTTCCACCACTTTCATTAAGATCCAG
GTGGAAGATGTTGATGAGCCTCCTCTTTTCTCCTTCCATATTATGTATTTGAAGTTTTTGA
AGAAACCCACAGGGATCATTTGTAGGCGTGGTGTCTGCCACAGACCCAGACAATAGGAAAT
CTCCTATCAGGTATTCTATTACTAGGAGCAAAGTGTTCAATATCAATGATAATGGTACAATC
ACTACAAGTAACTCACTGGATCGTGAAATCAGTGCTTGGTACAACCTAAGTATTACAGCCAC
AGAAAAATACAATATAGAACAGATCTCTTCGATCCCACCTGTATGTGCAAGTTCTTAACATCA
ATGATCATGCTCCTGAGTTCTCTCAATACTATGAGACTTATGTTTTGTGAAAATGCAGGCTCT
GGTCAGGTAATTCAGACTATCAGTGCAGTGGATAGAGATGAATCCATAGAAGAGCACCATTT
TTACTTTAATCTATCTGTAGAAGACACTAACAATTCAGTTTTTACAATCATAGATAATCAAG
ATAACACAGCTGTCATTTTGACTAATAGAACTGGTTTTTAACCTTCAAGAAGAACCTGTCTTC
TACATCTCCATCTTAATTGCCGACAATGGAATCCCGTCACTTACAAGTACAAACACCCCTTAC
CATCCATGTCTGTGACTGTGGTGACAGTGGGAGCACACAGACCTGCCAGTACCAGGAGCTTG
TGCTTTCCATGGGATTC AAGACAGAAGTTATCATTGCTATTCTCATTTCATTATGATCATA
TTTGGGTTTTATTTTTTTGACTTTGGGTTTAAAAACAACGGAGAAAACAGATTCTATTTCTGA
GAAAAGTGAAGATTT CAGAGAGAATATATTCCAATATGATGATGAAGGGGGTGGAGAAGAAG
ATACAGAGGCCTTTGATATAGCAGAGCTGAGGAGTAGTACCATAATGCGGGAACGCAAGACT
CGGAAAACCACAAGCGCTGAGATCAGGAGCCTATACAGGCAGTCTTTGCAAGTTGGCCCCGA
CAGTGCCATATT CAGGAAATTCATTCTGGAAGAGCTCGAAGAAGCTAATACTGATCCGTGTG
CCCCCTCCTTTTGATTCCCTCCAGACCTACGCTTTTGAGGGAACAGGGTCATTAGCTGGATCC
CTGAGCTCCTTAGAATCAGCAGTCTCTGATCAGGATGAAAGCTATGATTACCTTAATGAGTT
GGGACCTCGCTTTAAAAGATTAGCATGCATGTTTGGTTCTGCAGTGCAGTCAAATAATTAGG
GCTTTTTTACCATCAAATTTTTTAAAAGTGCTAATGTGTATT CGAACCCAATGGTAGTCTTAA
AGAGTTTTGTGCCCTGGCTCTATGGCGGGGAAAGCCCTAGTCTATGGAGTTTTCTGATTTCC
CTGGAGTAAATACTCCATGGTTATTTTTAAGCTACCTACATGCTGT CATTGAACAGAGATGTG
GGGAGAAATGTAAACAATCAGCTCACAGGCATCAATACAACCAGATTTGAAGTAAAATAATG
TAGGAAGATATTTAAAAGTAGATGAGAGGACACAAGATGTAGTCGATCCTTATGCGATTATAT
CAATTATTTACTTTAGGAAAGAGTAAAAATACCAAACGAGAAAATTTAAAGGAGCAAAAATTTG
CAAGTCAAATAGAAATGTACAAATCGAGATAACATTTACATTTCTATCATATTGACATGAAA
ATTGAAAATGTATAGTCAGAGAAATTTTCATGAATTATTCCATGAAGTATTGTTTTCTTTTAT
TTAAA

Year	Age	Sex	Occupation	Education	Marital Status	Religion	Political Party	Income	Health	Smoking	Alcohol	Exercise	Stress	Family Size	Community	Environment	Genetics	Other
1990	25	Male	Engineer	High School	Married	Catholic	Democrat	\$45,000	Good	Yes	No	Regular	Low	3	Suburban	Urban	None	
1991	26	Female	Teacher	College	Single	Protestant	Republican	\$38,000	Fair	No	No	Occasional	Medium	2	Rural	Rural	None	
1992	27	Male	Doctor	University	Married	Jewish	Democrat	\$75,000	Excellent	No	No	Daily	Low	4	Urban	Urban	None	
1993	28	Female	Nurse	College	Married	Muslim	Democrat	\$42,000	Good	Yes	No	Regular	Low	3	Suburban	Urban	None	
1994	29	Male	Lawyer	University	Single	Christian	Republican	\$65,000	Fair	No	No	Occasional	Medium	2	Rural	Rural	None	
1995	30	Female	Artist	College	Single	Buddhist	Democrat	\$35,000	Good	Yes	No	Regular	Low	3	Suburban	Urban	None	
1996	31	Male	Scientist	University	Married	Hindu	Democrat	\$55,000	Excellent	No	No	Daily	Low	4	Urban	Urban	None	
1997	32	Female	Manager	College	Married	Sikh	Democrat	\$48,000	Fair	No	No	Occasional	Medium	2	Rural	Rural	None	
1998	33	Male	Writer	University	Single	Christian	Republican	\$40,000	Good	Yes	No	Regular	Low	3	Suburban	Urban	None	
1999	34	Female	Engineer	College	Married	Muslim	Democrat	\$50,000	Excellent	No	No	Daily	Low	4	Urban	Urban	None	
2000	35	Male	Teacher	High School	Married	Jewish	Democrat	\$43,000	Fair	No	No	Occasional	Medium	2	Rural	Rural	None	
2001	36	Female	Nurse	College	Single	Protestant	Republican	\$39,000	Good	Yes	No	Regular	Low	3	Suburban	Urban	None	
2002	37	Male	Doctor	University	Married	Catholic	Democrat	\$76,000	Excellent	No	No	Daily	Low	4	Urban	Urban	None	
2003	38	Female	Lawyer	College	Married	Muslim	Democrat	\$66,000	Fair	No	No	Occasional	Medium	2	Rural	Rural	None	
2004	39	Male	Artist	University	Single	Buddhist	Democrat	\$36,000	Good	Yes	No	Regular	Low	3	Suburban	Urban	None	
2005	40	Female	Scientist	College	Married	Hindu	Democrat	\$56,000	Excellent	No	No	Daily	Low	4	Urban	Urban	None	
2006	41	Male	Manager	High School	Married	Sikh	Democrat	\$49,000	Fair	No	No	Occasional	Medium	2	Rural	Rural	None	
2007	42	Female	Writer	College	Single	Christian	Republican	\$41,000	Good	Yes	No	Regular	Low	3	Suburban	Urban	None	
2008	43	Male	Engineer	University	Married	Muslim	Democrat	\$51,000	Excellent	No	No	Daily	Low	4	Urban	Urban	None	
2009	44	Female	Teacher	College	Married	Jewish	Democrat	\$44,000	Fair	No	No	Occasional	Medium	2	Rural	Rural	None	
2010	45	Male	Nurse	High School	Married	Protestant	Republican	\$40,000	Good	Yes	No	Regular	Low	3	Suburban	Urban	None	
2011	46	Female	Doctor	University	Single	Catholic	Democrat	\$77,000	Excellent	No	No	Daily	Low	4	Urban	Urban	None	
2012	47	Male	Lawyer	College	Married	Muslim	Democrat	\$67,000	Fair	No	No	Occasional	Medium	2	Rural	Rural	None	
2013	48	Female	Artist	University	Single	Buddhist	Democrat	\$37,000	Good	Yes	No	Regular	Low	3	Suburban	Urban	None	
2014	49	Male	Scientist	College	Married	Hindu	Democrat	\$57,000	Excellent	No	No	Daily	Low	4	Urban	Urban	None	
2015	50	Female	Manager	High School	Married	Sikh	Democrat	\$50,000	Fair	No	No	Occasional	Medium	2	Rural	Rural	None	
2016	51	Male	Writer	College	Single	Christian	Republican	\$42,000										

><MW: 87002, pI: 4.64, NX(S/T): 8

amino acids 136-146 and 244-254

Year	Age	Sex	Location	Species	Number
1990	10	Male	1000	1000	1000
1991	10	Male	1000	1000	1000
1992	10	Male	1000	1000	1000
1993	10	Male	1000	1000	1000
1994	10	Male	1000	1000	1000
1995	10	Male	1000	1000	1000
1996	10	Male	1000	1000	1000
1997	10	Male	1000	1000	1000
1998	10	Male	1000	1000	1000
1999	10	Male	1000	1000	1000
2000	10	Male	1000	1000	1000
2001	10	Male	1000	1000	1000
2002	10	Male	1000	1000	1000
2003	10	Male	1000	1000	1000
2004	10	Male	1000	1000	1000
2005	10	Male	1000	1000	1000
2006	10	Male	1000	1000	1000
2007	10	Male	1000	1000	1000
2008	10	Male	1000	1000	1000
2009	10	Male	1000	1000	1000
2010	10	Male	1000	1000	1000
2011	10	Male	1000	1000	1000
2012	10	Male	1000	1000	1000
2013	10	Male	1000	1000	1000
2014	10	Male	1000	1000	1000
2015	10	Male	1000	1000	1000
2016	10	Male	1000	1000	1000
2017	10	Male	1000	1000	1000
2018	10	Male	1000	1000	1000
2019	10	Male	1000	1000	1000
2020	10	Male	1000	1000	1000

ATTTCAAGGCCAGCCATATTTTTNTGTTGAACCAACAACAGGAGTCATAAGAATATTTTNTA
 AAATGGATAGAGAACTGCAAGATGAGTATTGGGTAATCATTCAAGCCAAGGACATGATTGGT
 CAGCCAGGAGCGTTGTNTGGAACAACAAGTGTATTAATTAACTTTTCAGATGTTAATGACAA
 TAAGCCTATATTTAAAGAAAGTTTATACCGCTTGACTGTNTNTGAATCTGCACCCACTGGGA
 NTTNTATAGGAACAATCATGGCATATGATAATGACATAGGAGAGAATGCAGAAATGGATTAC
 AGCATTGAAGAGGATGATTTCGCAACATTTGACATTATT

FIGURE 97

GCAACCTCAGCTTCTAGTATCCAGACTCCAGCGCCGCCCCGGGCGCGGACCCCAACCCCGAC
CCAGAGCTTCTCCAGCGGCGGCGCAGCGAGCAGGGCTCCCCGCCTTAACCTCCTCCGCGGGG
CCCAGCCACCTTCGGGAGTCCGGGTTGCCCACCTGCAAACCTCTCCGCCTTCTGCACCTGCCA
CCCCTGAGCCAGCGCGGGCCCCCGAGCGAGTCAATGGCCAACGCGGGGCTGCAGCTGTTGGGC
TTCATTCTCGCCTTCTGCGGATGGATCGGCGCCATCGTCAGCACTGCCCTGCCCCAGTGGAG
GATTTACTCCTATGCCGCGCACAACATCGTGACCGCCAGGCCATGTACGAGGGGCTGTGGA
TGTCTGCGTGTFCGAGAGCACCGGGCAGATCCAGTGCAAAGTCTTTGACTCCTTGCTGAAT
CTGAGCAGCACATTGCAAGCAACCCGTGCCCTTGATGGTGGTTGGCATCCTCCTGGGAGTGAT
AGCAATCTTTGTGGCCACCGTTGGCATGAAGTGTATGAAGTGCTTGGGAAGACGATGAGGTGC
AGAAGATGAGGATGGCTGTCAATTGGGGGTGCGATATTTCTTCTTGACAGTCTGGCTATTTTA
GTTGCCACAGCATGGTATGGCAATAGAATCGTTCAAGAATTCTATGACCCTATGACCCAGT
CAATGCCAGGTACGAATTTGGTCAAGGCTCTCTTCACTGGCTGGGCTGCTGCTTCTCTCGCC
TTCTGGGAGGTGCCCTACTTTTGTGTTCTGTCCCCGAAAAACAACCTCTTACCCAACACCA
AGGCCCTATCCAAAACCTGCACCTTCCAGCGGGAAAGACTACGTGTGACACAGAGGCAAAAG
GAGAAAATCATGTTGAAACAACCGAAAATGGACATTGAGATACTATCATTAACATTAGGAC
CTTAGAATTTTGGGTATTGTAATCTGAAGTATGGTATTACAAAACAAACAAACAAAAA
ACCCATGTGTTAAATACTCAGTGCTAAACATGGCTTAATCTTATTTTATCTTCTTTCCTCA
ATATAGGAGGGAAGATTTTTCATTTGTATTACTGCTTCCCATTGAGTAATCATACTCAAAT
GGGGGAAGGGGTGCTCCTTAAATATATATAGATATGTATATATACATGTTTTTCTATTAAAA
ATAGACAGTAAATACTATTCTCATTATGTTGATACTAGCATACTTAAATATCTCTAAAAT
AGGTAAATGTATTTAATTCCATATTGATGAAGATGTTTATTGGTATATTTTCTTTTTCGTCC
TTATATACATATGTAACAGTCAAATATCATTTACTCTTCTTCATTAGCTTTGGGTGCCTTTG
CCACAAGACCTAGCCTAATTTACCAAGGATGAATTCTTTCAATTCTTCATGCGTGCCCTTTT
CATATACTTATTTTATTTTACCATAATCTTATAGCACTTGCATCGTTATTAAGCCCTTAT
TTGTTTTGTGTTTCATTGGTCTCTATCTCCTGAATCTAACACATTTTCATAGCCTACATTTTA
GTTTCTAAAGCCAAGAAGAATTTATTACAAATCAGAACTTTGGAGGCAAATCTTCTGCATG
ACCAAAGTGATAAATTCCTGTTGACCTTCCCACACAATCCCTGTACTCTGACCCATAGCACT
CTTGTTTGCTTTGAAAATATTTGTCCAATTGAGTAGCTGCATGCTGTTCCCCCAGGTGTTGT
AACACAACCTTTATTGATTGAATTTTAAAGCTACTTATTTCATAGTTTTATATCCCCCTAACT
ACCTTTTTTGTTCCCCATTCTTAATTGTATTGTTTTCCCAAGTGTAATTATCATGCGTTTTA
TATCTTCCTAATAAGGTGTGGTCTGTTTTGTCTGAACAAAGTGCTAGACTTTCTGGAGTGATA
ATCTGGTGACAAATATTCTCTCTGTAGCTGTAAGCAAGTCACTTAATCTTCTACCTCTTTT
TTCTATCTGCCAAATTGAGATAATGATACTTAACCAGTTAGAAGAGGTAGTGTGAATATTAA
TTAGTTTATATTACTCTTATTCTTTGAACATGAACTATGCCTATGTAGTGTCTTTATTTGCT
CAGCTGGCTGAGACACTGAAGAAGTCACTGAACAAAACCTACACACGTACCTTCATGTGATT
CACTGCCTTCCTCTCTCTACCAGTCTATTTCCACTGAACAAAACCTACACACATACCTTCAT
GTGGTTCAGTGCCTTCCTCTCTCTACCAGTCTATTTCCACTGAACAAAACCTACGCACATAC
CTTCATGTGGCTCAGTGCCTTCCTCTCTCTACCAGTCTATTTCCATTCTTTCAGCTGTGTCT
GACATGTTTGTGCTCTGTTCCATTTTAAACAACCTGCTCTTACTTTTCCAGTCTGTACAGAATG
CTATTTCACTTGAGCAAGATGATGTAATGGAAAGGGTGTGGCACTGGTGTCTGGAGACCTG
GATTTGAGTCTTGGTGCTATCAATCACCGTCTGTGTTTGGAGCAAGGCATTTGGCTGCTGTAA
GCTTATTGCTTCATCTGTAAGCGGTGGTTTGTAAATTCCTGATCTTCCCACCTCACAGTGATG
TTGTGGGGATCCAGTGAGATAGAATACATGTAAGTGTGGTTTTGTAAATTTAAAAAGTGCTAT
ACTAAGGGAAAGAATTGAGGAATTAAGTGCATACGTTTTGGTGTGCTTTTCAAATGTTTGA
AAATAAAAAAATGTTAAG

Year	Age	Sex	Location	Case No.	Ref.
1970	10-14	M	India	1	[1]
1971	15-19	F	India	2	[2]
1972	20-24	M	India	3	[3]
1973	25-29	F	India	4	[4]
1974	30-34	M	India	5	[5]
1975	35-39	F	India	6	[6]
1976	40-44	M	India	7	[7]
1977	45-49	F	India	8	[8]
1978	50-54	M	India	9	[9]
1979	55-59	F	India	10	[10]
1980	60-64	M	India	11	[11]
1981	65-69	F	India	12	[12]
1982	70-74	M	India	13	[13]
1983	75-79	F	India	14	[14]
1984	80-84	M	India	15	[15]
1985	85-89	F	India	16	[16]
1986	90-94	M	India	17	[17]
1987	95-99	F	India	18	[18]
1988	100+	M	India	19	[19]
1989	101+	F	India	20	[20]
1990	102+	M	India	21	[21]
1991	103+	F	India	22	[22]
1992	104+	M	India	23	[23]
1993	105+	F	India	24	[24]
1994	106+	M	India	25	[25]
1995	107+	F	India	26	[26]
1996	108+	M	India	27	[27]
1997	109+	F	India	28	[28]
1998	110+	M	India	29	[29]
1999	111+	F	India	30	[30]
2000	112+	M	India	31	[31]
2001	113+	F	India	32	[32]
2002	114+	M	India	33	[33]
2003	115+	F	India	34	[34]
2004	116+	M	India	35	[35]
2005	117+	F	India	36	[36]
2006	118+	M	India	37	[37]
2007	119+	F	India	38	[38]
2008	120+	M	India	39	[39]
2009	121+	F	India	40	[40]
2010	122+	M	India	41	[41]
2011	123+	F	India	42	[42]
2012	124+	M	India	43	[43]
2013	125+	F	India	44	[44]
2014	126+	M	India	45	[45]
2015	127+	F	India	46	[46]
2016	128+	M	India	47	[47]
2017	129+	F	India	48	[48]
2018	130+	M	India	49	[49]
2019	131+	F	India	50	[50]
2020	132+	M	India	51	[51]
2021	133+	F	India	52	[52]
2022	134+	M	India	53	[53]
2023	135+	F	India	54	[54]
2024	136+	M	India	55	[55]
2025	137+	F	India	56	[56]
2026	138+	M	India	57	[57]
2027	139+	F	India	58	[58]
2028	140+	M	India	59	[59]
2029	141+	F	India	60	[60]
2030	142+	M	India	61	[61]
2031	143+	F	India	62	[62]
2032	144+	M	India	63	[63]
2033	145+	F	India	64	[64]
2034	146+	M	India	65	[65]
2035	147+	F	India	66	[66]
2036	148+	M	India	67	[67]
2037	149+	F	India	68	[68]
2038	150+	M	India	69	[69]
2039	151+	F	India	70	[70]
2040	152+	M	India	71	[71]
2041	153+	F	India	72	[72]
2042	154+	M	India	73	[73]

><MW: 22744, pI: 8.51, NX(S/T): 1

amino acids 119-133

FIGURE 99

TTCTGGCCAAACCCGGGGCTNCAGCTGTTGGGCTTCATCTCGCCTTCCTGGGATGGATCGGC
GCCATCNTCACACTGCCCTTCCCCAGTGGAGGATTTTACTCCCTATGCTGGCGACAACATCG
TGACCGCCCAGCCCATGTACGAGGGGCTGTGGATGTCCNGCGTGTCGCAGAGCACCGGGCAG
ATCCAGTGCAAAGTCTTTGACTCCTTGCTGAATCTGAGCAGCACATTGCAAGCAACCCGTGC
CTTGATGGTGGTTGGCATCCTCCTGGGAGTGATAGCAATCTTTGTGGCCACCGTTGGCATGA
AGTGTATGAAGTGCTTGGAAGACGATGAGGTGCAGAAGATGAGGATGGCTGTCATTGGGGGC
GCGATATTTCTTCTTGCAAGTCTGGCTATTTTAGTTGCCACAGCATGGTATGGCAATAGAAN
CNTTCAACANTTCTATGACCCTATGACCCAGTCAATGCCAGGTACGAATTTGGTCA
GGCTCTCTTCACTGGCTGGGCTGCTGCTTCTCTCTGCCTTCTGGGAGGTGCCCTACTTTGCT
GTTCCTGTCCC

FIGURE 100

ACCCTTGACCCAACGCGGCCCCCGACCGNTTCATGGCCAAACGCGGGNCTCCAGCTGTTGG
GCTTCATTCTCCCCTTCCTGGGATGGACCGGCGCCCATCNTCAGCACTGCCCTGCCCCAGTG
GAGGATTTACTCCTATNCCGGCNACAACATCGTGACCGCCCAGGCCNTGTACGAGGGGCTGT
GGATGTCCTGCGTGTTCGCAGAGCACCGGGCAGATCCAGTGCAAAGTCTTTGACTCCCTTGCT
GAATCTGAGCAGCACATTGCAAGCAACCCGTGCCTTGATGGTGGTTGGCATCCTCCTGGGAG
TGATAGCAATCTTNNTGGCCACCGTTGTNNNTGAAGTGTATGAAGTGCTTGGAAGACGATGA
GGTGCAGAAGATGAGGATGGCTGTCATTGGGGGCGCGATATTTCTTCTTGCAGGTCTGGCTA
TTTTAGTTGCCACAGCATGGTATGGCAATAGAATCGTTCAAGAATTCTATGACCCTATGACCGA

[illegible]

GGGCCCGACCATATTATCCAACCGGGNTCACTGTTGGCTCATCTCCCTCCTGGATGAANCGCGC
CATCNTCAGACTCCCTGCCCCATGGAGATTTNNCCTATGCTGGCGACAACATCNTGACCCCC
AGCCATGTACGAGGGGCTTTGAACGTCNGCGTGTGCGAGANACCGGGCAGATCCAGTGCAA
AGTCTTTGACTCCTTGCTGAATCTGNGCAGCACATTGCAGCAACCCNTGCCCTGATGGTGGT
TGGCATCCTCCTGGGAGTGATAGCAATCTTTGTGGCCACCGTTGGCATGAAGTGTATGAAGT
GCTTGGAAGACGATGAGGTGCAGAAGATGAGGATGGCTGTCAATTGGGGGCGCGATATTTCTT
CTTG CAGGTCTGGCTATTTNNNGTTGCCACAGCATGGTATGGCAATAGAATCGTTCAAGAAT
TCTATGACCCCTATGACCCCAAGTCAATGCCAGGTACGAATTTGGTCAGGCTCTCTTCACTGGC
TGGGCTGCTGCTTCTCTCTGCCTTCTGGGAGGTGCCCTACTTTGCTGTTTCTTGCGA

FIGURE 102

ATTCTCCCCTCCTGGATGGATCGCNCCACCGTCACATTGCCTTCCCCCANTGGAGGATTNAC
TCCTATGCTGGCGACAACATCGTGACCCCCCAGGCCATTTACCGAGGGGCTTTGGATGTCNT
GCNTGTGCGCAGAGCACCGGGCAGATCCCAGTGCAAAGTCTTTGACTCCTTGCTGAATCTGAG
CAGCACATTGCAAGCAACCCGTGCCTTGATGGGGTTGGCATCCTCCTGGGAGTGATAGCAAC
CTTTGTGGCCACCGTTGGCATGAAGTGTATGAAGTGCTTGGAAGACGATGAGGTGCCAGAAG
ATGAGGATGGCTGTCATTGGGGGCGCGATATTTCTTGTTGCAGGTCTGGCTATTTTAGTNGC
CACAGCATGGTATGGCAATAGANTNNTTCNNGNNNTCTATGACCCTATGACCCAGTCAATG
CCAGGTACGAATTTGGTCAGGCTCTCTTCACTGGCTGGGCTGCTGCTTCTCTCTGCCTTCTG
GGAGGTGCCCTACTTTGCTGTTCCCTGTCCC

FIGURE 103

AGAGCACCGGCAGATCCCAGTNCAAAGTCTTTGACCCTTGCTGAATCTGAGCAGCACATTNC
AAGCAACCCCTTGCCTTGAAGGTGGTTGNCAATCCCCCTGGGAGTGAATAGCAATCTTTGTG
GCCACCGTTGGCATGAAGTNTATGAAGTGCTTGGAAGACGATGAGGTGCAGAAGATGAGGAT
GGCTGTCATTGGGGGCGCGATATTTCTTCTTGCAAGTCTGGCTATTTTAGTNNCCACAGCAT
GGTATGGCAATAGNATNNTTCGNGGNTTCTATGACCCTATGACCCCAGTCAATGCCAGGTAC
GAATTTGGTCAGGCTCTCTTCACTGGCTGGGCTGCTGCTTCTCTCTGCCTTCTGGGAGGTGC
CCTACTTTGCTGTTCTCTGTCCTCCCGAA

FIGURE 104

AGCAATGCCCTGCCCCAGTGGAGGATTAATTCCTATGNTGGGGACAACATTGTGACNGCCC
AGGCCATGTACGGGGGGCTGTGGATGTCCTGCGTGTGCGAGAGCACCGGGCAGATCCAGTGC
AAAGTNTTTGACTCCTTGCTGAATTTGAGCAGCACATTGCAAGCAACCCGTGCCTTGATGGT
GGTTGGCATCTTCCTGGGAGTGATAGCAATCTTTGTGGCCACCGTGGNAATGAAGTGTATGA
AGTGCTTGGAAGACGATGAGGTGCAGAAGATGAGGATGGCTGTCATTGGGGGCGCGATATTT
CTTNTTGCAGGTCTGGCTATTTTAGTTGCCACAGCATGGTATGGCAATAGAATNGTTCAAGA
ATTTTATGACCCTATGACCCCAGTCAATGCCAGGTACGAATTTGGTCAGGCTTTNTTCACTG
GCTGGGCTGCTGCTTNTTTCTGCCTTNTGGGAGGTGCCCTANTTTGCTGTTCTGCGAACC

FIGURE 105

TCATAGGGGGGCGCGATATTTTTCTTGCAGGTNTGGTTATTTTAGTTGCCACAGCATGGTA
TGGCAATAGAATCGTTCAAGAATTNTATGACCCTATGACCCAGTCAATGCCAGGTACGAAT
TTGGTCAGGCTCTNTTCACTGGNTGGGCTGCTGCTTCTNTNNGCCTTNTGGGAGGTGCCCTA
CTTTGCTGTTTCCTG

4000367-103404

FIGURE 106

TTCCTGGGATGGATCCGCCCCATCNTCACATGCCCTGCCCCNTGGAGATTTACNCCTATGC
TGGCGAACAACATCNTGACCGCCCAGGCCATGTACGAGGGGCTGTGGAATGTCCTGCGTGTC
CCAGAGCACCGGGCAGATCCAGTGCAAAGTCTTTGACTCCTTGCTGAATCTGAGCAGCACAT
TGCAAGCAACCNTGCCTTGATGGTGGTTGGCATCCTCCTGGGAGTGATAGCAATCTTTGTGG
CCACCGTTGGCATGAAAGTGTATGAAGTGCTTGAAGACGATGAGGTGCAGAAGATGAGGAT
GGCTGTCATTGGGGGCGCGATATTTCTTCTTGCAAGTCTGGCTATTTTAGNNGCCACAGCAT
GGTATGGCAATCAGACCCNNTCANAACTCTATGACCCCTATGACCCAGTCAATGCCAGGTA
CGAATTTGGTCAGGCTCTCTTCACTGGCTGGGCTGCTGCTTCTCTCTGCCTTCTGGGAGGTG
CCCTACTTTGCTGTTCTGTCCCCGAAAAACAACCTCTTACCCACG

Variable	Mean	SD	Min	Max
Age	38.5	10.2	22	65
Gender	0.5	0.5	0	1
Marital status	0.7	0.5	0	1
Education	12.5	2.5	9	16
Income	15.5	10.5	5	45
Health status	1.5	0.5	1	2
Stress level	3.5	1.5	1	5
Life satisfaction	4.5	1.5	1	7
Work satisfaction	5.5	1.5	1	7
Family satisfaction	6.5	1.5	1	7
Community satisfaction	5.5	1.5	1	7
Overall satisfaction	5.5	1.5	1	7
Life expectancy	75.5	5.5	65	85
Quality of life	65.5	10.5	50	80
Healthcare access	7.5	1.5	5	10
Healthcare quality	6.5	1.5	4	10
Healthcare cost	8.5	1.5	6	10
Healthcare coverage	9.5	0.5	9	10
Healthcare satisfaction	7.5	1.5	5	10
Healthcare utilization	6.5	1.5	4	10
Healthcare expenditure	7.5	1.5	5	10
Healthcare equity	6.5	1.5	4	10
Healthcare efficiency	7.5	1.5	5	10
Healthcare effectiveness	6.5	1.5	4	10
Healthcare safety	7.5	1.5	5	10
Healthcare transparency	6.5	1.5	4	10
Healthcare accountability	7.5	1.5	5	10
Healthcare integrity	6.5	1.5	4	10
Healthcare honesty	7.5	1.5	5	10
Healthcare fairness	6.5	1.5	4	10
Healthcare justice	7.5	1.5	5	10
Healthcare freedom	6.5	1.5	4	10
Healthcare peace	7.5	1.5	5	10
Healthcare love	6.5	1.5	4	10
Healthcare kindness	7.5	1.5	5	10
Healthcare compassion	6.5	1.5	4	10
Healthcare empathy	7.5	1.5	5	10
Healthcare sympathy	6.5	1.5	4	10
Healthcare pity	7.5	1.5	5	10
Healthcare mercy	6.5	1.5	4	10
Healthcare grace	7.5	1.5	5	10
Healthcare forgiveness	6.5	1.5	4	10
Healthcare tolerance	7.5	1.5	5	10
Healthcare patience	6.5	1.5	4	10
Healthcare perseverance	7.5	1.5	5	10
Healthcare determination	6.5	1.5	4	10
Healthcare resolve	7.5	1.5	5	10
Healthcare courage	6.5	1.5	4	10
Healthcare bravery	7.5	1.5	5	10
Healthcare valor	6.5	1.5	4	10
Healthcare heroism	7.5	1.5	5	10
Healthcare gallantry	6.5	1.5	4	10
Healthcare magnanimity	7.5	1.5	5	10
Healthcare generosity	6.5	1.5	4	10
Healthcare largesse	7.5	1.5	5	10
Healthcare munificence	6.5	1.5	4	10
Healthcare beneficence	7.5	1.5	5	10
Healthcare philanthropy	6.5	1.5	4	10
Healthcare altruism	7.5	1.5	5	10
Healthcare selflessness	6.5	1.5	4	10
Healthcare unselfishness	7.5	1.5	5	10
Healthcare disinterestedness	6.5	1.5	4	10
Healthcare impartiality	7.5	1.5	5	10
Healthcare objectivity	6.5	1.5	4	10
Healthcare fairness	7.5	1.5	5	10
Healthcare justice	6.5	1.5	4	10
Healthcare equity	7.5	1.5	5	10
Healthcare balance	6.5	1.5	4	10
Healthcare harmony	7.5	1.5	5	10
Healthcare concord	6.5	1.5	4	10
Healthcare agreement	7.5	1.5	5	10
Healthcare consensus	6.5	1.5	4	10
Healthcare accord	7.5	1.5	5	10
Healthcare understanding	6.5	1.5	4	10
Healthcare comprehension	7.5	1.5	5	10
Healthcare knowledge	6.5	1.5	4	10
Healthcare wisdom	7.5	1.5	5	10
Healthcare insight	6.5	1.5	4	10
Healthcare intellect	7.5	1.5	5	10
Healthcare intelligence	6.5	1.5	4	10
Healthcare reason	7.5	1.5	5	10
Healthcare logic	6.5	1.5	4	10
Healthcare rationality				

Variable	Mean	SD	Min	Max
Age	38.5	10.2	22	65
Gender	0.5	0.5	0	1
Marital status	0.7	0.5	0	1
Education	12.5	2.5	9	16
Income	15.5	10.5	5	45
Health status	1.5	0.5	1	2
Stress level	3.5	1.5	1	5
Life satisfaction	4.5	1.5	1	7
Work satisfaction	5.5	1.5	1	7
Family satisfaction	6.5	1.5	1	7
Community satisfaction	5.5	1.5	1	7
Overall satisfaction	5.5	1.5	1	7
Life expectancy	75.5	5.5	65	85
Quality of life	65.5	10.5	50	80
Healthcare access	7.5	1.5	5	10
Healthcare quality	6.5	1.5	4	10
Healthcare cost	8.5	1.5	6	10
Healthcare coverage	9.5	0.5	9	10
Healthcare satisfaction	7.5	1.5	5	10
Healthcare utilization	6.5	1.5	4	10
Healthcare expenditure	7.5	1.5	5	10
Healthcare equity	6.5	1.5	4	10
Healthcare efficiency	7.5	1.5	5	10
Healthcare effectiveness	6.5	1.5	4	10
Healthcare safety	7.5	1.5	5	10
Healthcare transparency	6.5	1.5	4	10
Healthcare accountability	7.5	1.5	5	10
Healthcare integrity	6.5	1.5	4	10
Healthcare honesty	7.5	1.5	5	10
Healthcare fairness	6.5	1.5	4	10
Healthcare justice	7.5	1.5	5	10
Healthcare freedom	6.5	1.5	4	10
Healthcare peace	7.5	1.5	5	10
Healthcare love	6.5	1.5	4	10
Healthcare kindness	7.5	1.5	5	10
Healthcare compassion	6.5	1.5	4	10
Healthcare empathy	7.5	1.5	5	10
Healthcare sympathy	6.5	1.5	4	10
Healthcare pity	7.5	1.5	5	10
Healthcare mercy	6.5	1.5	4	10
Healthcare grace	7.5	1.5	5	10
Healthcare forgiveness	6.5	1.5	4	10
Healthcare tolerance	7.5	1.5	5	10
Healthcare patience	6.5	1.5	4	10
Healthcare perseverance	7.5	1.5	5	10
Healthcare determination	6.5	1.5	4	10
Healthcare resolve	7.5	1.5	5	10
Healthcare courage	6.5	1.5	4	10
Healthcare bravery	7.5	1.5	5	10
Healthcare valor	6.5	1.5	4	10
Healthcare heroism	7.5	1.5	5	10
Healthcare gallantry	6.5	1.5	4	10
Healthcare magnanimity	7.5	1.5	5	10
Healthcare generosity	6.5	1.5	4	10
Healthcare largesse	7.5	1.5	5	10
Healthcare munificence	6.5	1.5	4	10
Healthcare beneficence	7.5	1.5	5	10
Healthcare philanthropy	6.5	1.5	4	10
Healthcare altruism	7.5	1.5	5	10
Healthcare selflessness	6.5	1.5	4	10
Healthcare unselfishness	7.5	1.5	5	10
Healthcare disinterestedness	6.5	1.5	4	10
Healthcare impartiality	7.5	1.5	5	10
Healthcare objectivity	6.5	1.5	4	10
Healthcare fairness	7.5	1.5	5	10
Healthcare justice	6.5	1.5	4	10
Healthcare equity	7.5	1.5	5	10
Healthcare balance	6.5	1.5	4	10
Healthcare harmony	7.5	1.5	5	10
Healthcare concord	6.5	1.5	4	10
Healthcare agreement	7.5	1.5	5	10
Healthcare consensus	6.5	1.5	4	10
Healthcare accord	7.5	1.5	5	10
Healthcare understanding	6.5	1.5	4	10
Healthcare comprehension	7.5	1.5	5	10
Healthcare knowledge	6.5	1.5	4	10
Healthcare wisdom	7.5	1.5	5	10
Healthcare insight	6.5	1.5	4	10
Healthcare intellect	7.5	1.5	5	10
Healthcare intelligence	6.5	1.5	4	10
Healthcare reason	7.5	1.5	5	10
Healthcare logic	6.5	1.5	4	10
Healthcare rationality				

FIGURE 108

CGGTGCCGTCAGCTCGCCGGGCACCGCGGCCTCGCCCTCGCCCTCCGCCCCCTGCGCCTGCAC
CGCGTAGACCGACCCCCCCTCCAGCGCGCCACCCGGTAGAGGACCCCGCCCCGTGCCCCG
ACCGGTCCCCGCCTTTTTGTAAACTTTAAAGCGGGCGCAGCATTAAACGCTTCCCGCCCCGGT
GACCTCTCAGGGGTCTCCCCGCCAAAGGTGCTCCGCGCTAAGGAACATGGCGAAGGTGGAG
CAGGTCCTGAGCCTCGAGCCGCAGCACGAGCTCAAATTCCGAGGTCCCTTCACCGATGTTGT
CACCACCAACCTAAAGCTTGGCAACCCGACAGACCGAAATGTGTGTTTTAAGGTGAAGACTA
CAGCACCACGTAGGTACTGTGTGAGGCCCAACAGCGGAATCATCGATGCAGGGGCCTCAATT
AATGTATCTGTGATGTTACAGCCTTTTCGATTATGATCCCAATGAGAAAAGTAAACACAAGTT
TATGGTTTCAGTCTATGTTTTGCTCCAACCTGACACTTCAGATATGGAAGCAGTATGGAAGGAGG
CAAAACCGGAAGACCTTATGGATTCAAACTTAGATGTGTGTTTGAATTGCCAGCAGAGAAT
GATAAACCACATGATGTAGAAATAAAATAAAATTATATCCACAACCTGCATCAAAGACAGAAAC
ACCAATAGTGTCTAAGTCTCTGAGTTCCTTCTTTGGATGACACCGAAGTTAAGAAGGTTATGG
AAGAATGTAAGAGGCTGCAAGGTGAAGTTCAGAGGCTACGGGAGGAGAACAAGCAGTTCAAG
GAAGAAGATGGAATGCGGATGAGGAAGACAGTGCAGAGCAACAGCCCCATTTTCAGCATTAGC
CCCAACTGGGAAGGAAGAAGGCCTTAGCACCCGGCTCTTGGCTCTGGTGGTTTTTGTCTTTA
TCGTTGGTGTAATTATTGGGAAGATTGCCTTGTAGAGGTAGCATGCACAGGATGGTAAATTG
GATTGGTGGATCCACCATATCATGGGATTTAAATTTATCATAACCATGTGTAAAAAGAAATT
AATGTATGATGACATCTCACAGGTCTTGCCCTTTAAATTACCCCTCCCTGCACACACATACAC
AGATACACACACACAAATATAATGTAACGATCTTTTAGAAAGTTAAAAATGTATAGTAACTG
ATTGAGGGGGAAAAAGAATGATCTTTATTAATGACAAGGGAAACCATGAGTAATGCCACAAT
GGCATATTGTAAATGTCATTTTAAACATTGGTAGGCCTTGGTACATGATGCTGGATTACCTC
TCTTAAATGACACCTTTCCTCGCCTGTTGGTGCTGGCCCTTGGGGAGCTGGAGCCCAGCAT
GCTGGGGAGTGCGGTGAGCTCCACACAGTAGTCCCCACGTGGCCCACTCCCGGCCCAGGCTG
CTTTCCGTGTCTTCAGTTCTGTCCAAGCCATCAGCTCCTTGGGACTGATGAACAGAGTCAGA
AGCCCAAAGGAATTGCACTGTGGCAGCATCAGACGTACTCGTCATAAGTGAGAGGCGTGTGT
TGACTGATTGACCCAGCGCTTTGGAAATAAATGGCAGTGCTTTGTTCACTTAAAGGGACCAA
GCTAAATTTGTATTGGTTCATGTAGTGAAGTCAAACCTGTTATTTCAGAGATGTTTAAATGCATA
TTTAACTTATTTAATGTATTTTCATCTCATGTTTTCTTATTGTCAAGAGTACAGTTAATGC
TGCGTGCTGCTGAACTCTGTTGGGTGAACTGGTATTGCTGCTGGAGGGCTGTGGGCTCCTCT
GTCTCTGGAGAGTCTGGTCATGTGGAGGTGGGGTTTTATTGGGATGCTGGAGAAGAGCTGCCA
GGAAGTGTTTTTTCTGGGTGAGTAAATAACAACTGTTCATAGGGAGGGAAATTCTCAGTAGTG
ACAGTCAACTCTAGGTTACCTTTTTTAATGAAGAGTAGTCAGTCTTCTAGATTGTTCTTATA
CCACCTCTCAACCATTACTCACACTTCCAGCGCCAGGTCCAAGTCTGAGCCTGACCTCCCC
TTGGGGACCTAGCCTGGAGTCAGGACAAATGGATCGGGCTGCAGAGGGTTAGAAGCGAGGGC
ACCAGCAGTTGTGGGTGGGGAGCAAGGGAAGAGAGAACTCTTCAGCGAATCCTTCTAGTAC
TAGTTGAGAGTTTGACTGTGAATTAATTTTATGCCATAAAAGACCAACCCAGTTCTGTTTGA
CTATGTAGCATCTTGAAGAAATAATTAATAAAGCCCCAAAATTAAGAAAA

FIGURE 109

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA53977

<subunit 1 of 1, 243 aa, 1 stop

<MW: 27228, pI: 7.43, NX(S/T): 2

MAKVEQVLSLEPQHCLKFRGPFTDVVTTNLKLGNPTRNVCFKVKTTAPRRYCVRPNSGIID
AGASINVSVMQLPFDYDPNEKSKHKFMVQSMFAPTDTSMEAVWKEAKPEDLMDSKLRVFE
LPAENDKPHDVEINKIISTTASKTETPIVSKSLSSSLDDTEVKKVMEECKRLQGEVQRLREE
NKQFKEEDGLRMRKTVQSNPISALAPTGKEEGLSTRLLALVVLFFIVGVIIGKIAL

Important features:

Transmembrane domain:

amino acids 224-239

N-glycosylation site.

amino acids 68-71

N-myristoylation site.

amino acids 59-64, 64-69 and 235-240

FIGURE 110

GTCAGTCTTCTAGATTGTCCTTATCCCACCTTTCAACCANTACTCACATTTTCNAGCGCCCAG
GTCCANGTCTGAGCCTGACTTCCCCTTGGGGACCTAGCCTGGAGTCAGGACAATGGNTCGGG
CTGCAGAGGNTTAGAAGCGAGGGCACCAGCAGTTTTGGGTGGGGAGCAAGGGNNGAGAGAAA
CTCTTCAGCGAATCCTTCTAGTACTAGTTGAGAGTTTGACTGTGAATTAATTTTATGCCATA
AAAGACNAACCCAGTTCTGTTTGACTATGTAGCATCTTGAAAAGAAAAATTATAATAAGCC
CCAAAATTAAGAATTCTTTTGTCAATTTGTGCACATTTGCTCTATGGGGGGAATTATTATTTT
ATCATTTTTATTATTTTGCCATTGGAAGGTAACTTTAAAATGAGC

FIGURE 112

CCCTGGTGGTTTTGTTCTTTAATTCGTTGGTGTAATTNTTGGGAAGATTGCTTGTAGAGGTA
GNATGCACCNGGCTGGTAAATTGGATTGGTGGATCCACCATATCCATGGGATTTAAATTTAT
CATAACCATGTGTAAAAAGAAATTAATGTATGATGACATNTCACAGGTATTGCCTTTAAATT
ACCCATCCCTGNANACACATACACAGATACACANANACAAATNTAATGTAACGATNTTTTAG
AAAGTTAAAAATGTATAGTAAC

404637292000

[illegible]

GGTGGCCCATTCCCGGCCAGGCTGCTTTCGGTNTTCAGTTCCTGTCCAAGCCATCAGCTCC
TTGGGACTGATGAACAGAGTCAGAAGCCCAAAGGAATTGCACTGTGGCAGCATNAGACGTAC
TTGTNATAAGTGAGAGGCGTGTGTTGACTGATTGACCCAGCGCTTTTGAAATAAATGGCAGT
GCTTTGTTCANTTAAAGGGACCAAGCTAAATTTGTATTGGTTCATGTAGTGAAGTCAAAGT
TTATTTCAGAGATGTTTAATGCATATTTAANTTATTTAATGTATTTNATNTCATGTTTTCTTA
TTGTCACAAGAGTACAGTTAATGCTGCGTGCTGCTGAANTNTGTTGGGTGAACTGGTATTGC
TGCTGGAGGGCTGTGGGCTCCTCTGTCTTTGGAGAGTCTGGTCATGTGGAGGTGGG

[illegible][illegible]

FIGURE 115

AAACCTTTAAAAGTTGAGGGGAAAAGAATGATCCTTTATTAATGACAAGGGAAACNTGNGT
AATGCCACAATGGCATATTGTAAATGTCATTTTAAACATTGGTAGGCCTTGGTACATGATGC
TGGATTACCTCTCTTAAAATGACACCCTTCCTCGCCTGTTGGTGCTGGCCCTTGGGGAGCTN
GAGCCCAGCATGCTGGGGAGTGCGGTCTGCTCCACACAGTAGTCCCCANGTGGCCCANTCCC
GGCCCAGGCTGCTTTCCGTGTCTTCAGTTCTGTCCAAGCCATCAGCTCCTTGGGANTGATGA
ACAGAGTCAGAAGCCCAAAGGAATTGCANTGTGGCAGCATCAGANGTANTNGTCATAAGTGA
GAGGCGTGTGTTGANTGATTGACCCAGCGCTTTGGAAATAAATGGCAGTGCTTTGTTcantt
AAAGGGNCCAAGNTAAATTTGTATTGGTTCATGTAGTGAAGTCAAANTGTTATTcagagatg
TTTAATGCATATTTAANTTATTTAATGTATTTcATNTCATGTTTTCTTATTGTCACAAGGGT
ACAGTTAATGCTGCGTGCTGCTGAANTCTGTTGGGTGAANTGGTATTGCTG

[illegible]

GGCCCTTGGGGAGCTGGAGCCCAGCATGCTGGGGAGTGCGGTCAGCTCCACACAGTAGTCCC
CACGTGGCCCACTCCCGGCCAGGCTGCTTTCCGTGTCTTCAGTTCTGTCCAAGCCATCAGC
TCCTTGGGACTGATGAACAGAGTCAGAAGCCCAAAGGAATTGCACTGTGGCAGCATCAGACG
TACTCGTCATAAGTGAGAGGCGTGTGTTGACTGATTGACCCAGCGCTTTGGAAATAAATGGC
AGTGCTTTGTTCACTTAAAGGGACCAAGCTAAATTTGTATTGGTTCATGTAGTGAAGTCAA
CTGTTATTTCAGAGATGTTTAATGCATATTTAACTTATTTAATGTATTTTCATCTCATGTTTTTC
TTATTGTCAAGAGTACAGTTAATGCTGCGTGCTGCTGAACTCTGTTGGGTGAACTGGTAT
TGCTGCTGGAGGGCTGTGGGCTCCTCTGTCTCTGGAGAGTCTGGTCATGTGGAGGTGGG

FIGURE 117

GCAGCTCCGGGTGCTGTGGCCCCGGCCTTGGCGGGGCGGCCTCCGGCTCAGGCTGGCTGAGA
GGCTCCCAGCTGCAGCGTCCCCGCCCGCCTCCTCGGGAGCTCTGATCTCAGCTGACAGTGCC
CTCGGGGACCAAACAAGCCTGGCAGGGTCTCACTTTGTTGCCAGGCTGGAGTTCAGTGCCA
TGATCATGGTTTACTGCAGCCTTGACCTCCTGGGTTCAAGCGATCCTGCTGAGTAGCTGGGA
CTACAGGACAAAATTAGAAGATCAAAATGGAAAATATGCTGCTTTGGTTGATATTTTTCACC
CCTGGGTGGACCCTCATTGATGGATCTGAAATGGAATGGGATTTTATGTGGCACTTGAGAAA
GGTACCCCGGATTGTGAGTGAAGGACTTTCCATCTCACCAGCCCCGCATTTGAGGCAGATG
CTAAGATGATGGTAAATACAGTGTGTGGCATCGAATGCCAGAAAGAACTCCCAACTCCCAGC
CTTTCTGAATTGGAGGATTATCTTTCCTATGAGACTGTCTTTGAGAATGGCACCCGAACCTT
AACCAGGGTGAAAGTTCAAGATTTGGTTCTTGAGCCGACTCAAAATATCACCACAAAGGGAG
TATCTGTTAGGAGAAAGAGACAGGTGTATGGCACCGACAGCAGGTTGAGCATCTTGGACAAA
AGGTTCTTAACCAATTTCCCTTTCAGCACAGCTGTGAAGCTTTCCACGGGCTGTAGTGGCAT
TCTCATTTCCCCTCAGCATGTTCTAACTGCTGCCCACTGTGTTGATGAGAAAGGACTATG
TCAAAGGGAGTAAAAAGCTAAGGGTAGGGTTGTTGAAGATGAGGAATAAAAGTGGAGGCAAG
AAACGTCGAGGTTCTAAGAGGAGCAGGAGAGAAGCTAGTGGTGGTGACCAAAGAGAGGGTAC
CAGAGAGCATCTGCAGGAGAGAGCGAAGGGTGGGAGAAGAAGAAAAAATCTGGCCGGGGTC
AGAGGATTGCCGAAGGGAGGCCTTCTTTTCACTGGACCCGGGTCAAGAATACCCACATTCCG
AAGGGCTGGGCACGAGGAGGCATGGGGGACGCTACCTTGGAATGACTATGCTCTTCTGGA
GCTGAAGCGTGCTCACAAAAAGAAATACATGGAACCTTGAATCAGCCCAACGATCAAGAAAA
TGCCTGGTGAATGATCCACTTCTCAGGATTTGATAACGATAGGGCTGATCAGTTGGTCTAT
CGGTTTTGAGTGTGTCCGACGAATCCAATGATCTCCTTTACCAATACTGCGATGCTGAGTC
GGGCTCCACCGGTTCTGGGGTCTATCTGCGTCTGAAAGATCCAGACAAAAAGAATTGGAAGC
GCAAAATCATTGCGGTCTACTCAGGGCACCAGTGGGTGGATGTCCACGGGGTTGAGAAGGAC
TACAACGTTGCTGTTTCGCATCACTCCCCTAAAATACGCCAGATTTGCCTCTGGATTACCGG
GAACGATGCCAATTGTGCTTACGGCTAACAGAGACCTGAAACAGGGCGGTGTATCATCTAAA
TCACAGAGAAAACCAGCTCTGCTTACCGTAGTGAGATCACTTCATAGGTTATGCCTGGACTT
GAACTCTGTCAATAGCATTTCAACATTTTCAAATCAGGAGATTTTCGTCCATTTAAAAAA
TGTATAGGTGCAGATATTGAAACTAGGTGGGCACTTCAATGCCAAGTATATACTCTTCTTTA
CATGGTGATGAGTTTCATTTGTAGAAAAATTTTGTGCTTCTTAAAAATTAGACACACTTT
AAACCTTCAAACAGGTATTATAAATAACATGTGACTCCTTAATGGACTTATTCTCAGGGTCC
TACTCTAAGAAGAATCTAATAGGATGCTGGTTGTGTATTAAATGTGAAATTGCATAGATAAA
GGTAGATGGTAAAGCAATTAGTATCAGAATAGAGACAGAAAGTTACAACACAGTTTGTACTA
CTCTGAGATGGATCCATTGAGCTCATGCCCTCAATGTTTATATTGTGTTATCTGTTGGGTCT
GGGACATTTAGTTTAGTTTTTTTTGAAGAATTACAAATCAGAAGAAAAAGCAAGCATTATAAA
CAAACTAATAACTGTTTTACTGCTTTAAGAAATAACAATTACAATGTGTATTATTTAAAAA
TGGGAGAAATAGTTTGTCTATGAAATAAACCTAGTTTAGAAATAGGGAAGCTGAGACATTT
TAAGATCTCAAGTTTTTTATTTAACTAATACTCAAAATATGGACTTTTCATGTATGCATAGGG
AAGACACTTCACAAATTATGAATGATCATGTGTTGAAAGCCACATTATTTTATGCTATACAT
TCTATGTATGAGGTGCTACATTTTTAGGACAAAGAATTCTGTAATCTTTTTCAAGAAAGAGT
CTTTTTCTCCTTGACAAAATCCAGCTTTTGTATGAGGACTATAGGGTGAATTCTCTGATTAG
TAATTTTAGATATGTCCTTTCCTAAAAATGAATAAAATTTATGAATATGA

FIGURE 118

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA57253

<subunit 1 of 1, 413 aa, 1 stop

<MW: 47070, pI: 9.92, NX(S/T): 3

MENMLLWLIFFTPGWTLIDGSEMEWDFMWHLRKVPRIVSERTFHLTSPAFEADAKMMVNTVC
GIECQKELPTPSLSELEDYLSYETVFENGTRTLTRVKVQDLVLEPTQNITTKGVSVRRKRQV
YGTDSRFSILDKRFLTNFPFSTAVKLSTGCSGILISPQHVLTAAHCVHDGKDYVKGSKKLRV
GLLKMRNKS GGGKKRRGSKRSRREASGGDQREGTREHLQERAKGGRRRKSGRGQRIAEGRPS
FQWTRVKNTHIPKGWARGMGMDATLDYDYALLELKRAHKKKYMELGISPTIKKMPGGMIHFS
GFDNDRADQLVYRFCSVSDSNLLYQYCDAESGSTGSGVYLRLLKDPDKKNWKRKIIAVYSG
HQWVDVHGVQKDYNAVAVRITPLKYAQICLWIHGNDANCAYG

Important features:

Signal peptide:

amino acids 1-16

N-glycosylation sites.

amino acids 90-93, 110-113 and 193-196

Glycosaminoglycan attachment site.

amino acids 236-239

Serine proteases, trypsin family, histidine active site.

amino acids 165-170

FIGURE 119

AATGTGAGAGGGGCTGATGGAAGCTGATAGGCAGGACTGGAGTGTTAGCACCAGTACTGGAT
GTGACAGCAGGCAGAGGAGCACTTAGCAGCTTATTAGTGTCCGATTCTGATTCCGGCAAGG
ATCCAAGCATGGAATGCTGCCGTCGGGCAACTCCTGGCACACTGCTCCTCTTTCTGGCTTTC
CTGCTCCTGAGTTCCAGGACCGCACGCTCCGAGGAGGACCGGGACGGCCTATGGGATGCCTG
GGGCCCATGGAGTGAATGCTCACGCACCTGCGGGGGAGGGGCCTCCTACTCTCTGAGGCGCT
GCCTGAGCAGCAAGAGCTGTGAAGGAAGAAATATCCGATACAGAACATGCAGTAATGTGGAC
TGCCCACCAGAAGCAGGTGATTTCCGAGCTCAGCAATGCTCAGCTCATAATGATGTCAAGCA
CCATGGCCAGTTTTATGAATGGCTTCCTGTGTCTAATGACCCTGACAACCCATGTTCACTCA
AGTGCCAAGCCAAAGGAACAACCCTGGTTGTTGAACTAGCACCTAAGGTCTTAGATGGTACG
CGTTGCTATACAGAATCTTTGGATATGTGCATCAGTGCTTTATGCCAAATTGTTGGCTGCGA
TCACCAGCTGGGAAGCACCGTCAAGGAAGATAACTGTGGGGTCTGCAACGGAGATGGGTCCA
CCTGCCGGCTGGTCCGAGGGCAGTATAAATCCCAGCTCTCCGCAACCAAATCGGATGATACT
GTGGTTGCACTTCCCTATGGAAGTAGACATATTCGCCTTGTCTTAAAGGTCTCTGATCACTT
ATATCTGGAAACCAAACCCTCCAGGGGACTAAAGGTGAAAACAGTCTCAGCTCCACAGGAA
CTTTCCTTGTGGACAATTCTAGTGTGGACTTCCAGAAATTTCCAGACAAAGAGATACTGAGA
ATGGCTGGACCACTCACAGCAGATTTCAATTGTCAAGATTCGTAACCTCGGGCTCCGCTGACAG
TACAGTCCAGTTCATCTTCTATCAACCCATCATCCACCGATGGAGGGAGACGGATTTCTTTC
CTTGCTCAGCAACCTGTGGAGGAGTTATCAGCTGACATCGGCTGAGTGCTACGATCTGAGG
AGCAACCGTGTGGTTGCTGACCAATACTGTCACTATTACCCAGAGAACATCAAACCCAAACC
CAAGCTTCAGGAGTGCAACTTGGATCCTTGTCCAGCCAGTGACGGATACAAGCAGATCATGC
CTTATGACCTCTACCATCCCCTTCCTCGGTGGGAGGCCACCCCATGGACCGCGTGCTCCTCC
TCGTGTGGGGGGGGCATCCAGAGCCGGGCAGTTTCCTGTGTGGAGGAGGACATCCAGGGGCA
TGTCACCTTCAGTGGAAGAGTGGAATGCATGTACACCCCTAAGATGCCCATCGCGCAGCCCT
GCAACATTTTTTGA CTGCCCTAAATGGCTGGCACAGGAGTGGTCTCCGTGCACAGTGACATGT
GGCCAGGGCCTCAGATACCGTGTGGTCCTCTGCATCGACCATCGAGGAATGCACACAGGAGG
CTGTAGCCCAAAAACAAAGCCCCACATAAAAGAGGAATGCATCGTACCCACTCCCTGCTATA
AACCCAAAGAGAAACTTCCAGTCGAGGCCAAGTTGCCATGGTTCAAACAAGCTCAAGAGCTA
GAAGAAGGAGCTGCTGTGTCAGAGGAGCCCTCGTTAAGTTGTAAAAGCACAGACTGTTCTATA
TTTGAAACTGTTTTGTTTAAAGAAAGCAGTGTCTCACTGGTTGTAGCTTTCATGGGTTCTGA
ACTAAGTGTAATCATCTCACCAAAGCTTTTTGGCTCTCAAATTAAAGATTGATTAGTTTCAA
AAAAAAAAA

FIGURE 120

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA58847

<subunit 1 of 1, 525 aa, 1 stop

<MW: 58416, pI: 6.62, NX(S/T): 1

MECCRRATPGTLLLFLAFLLLSSRTARSEEDRDGLWDAWGPWSECSRTC GGGASYSLRRCLS
SKSCEGRNIRYRTC SNVDCPPEAGDFRAQQCSAHNDVKHHGQFYEWLPVSNPDNPSCLKCQ
AKGTTLVVELAPKVL DGTRCYTESLDMCISGLCQIVGCDHQLGSTVKEDNCGVCNGDGSTCR
LVRGQYKSQLSATKSDDTVVALPYGSRHIRLV LKGPDLHLYLETKTLQGTKGENSLSSSTGTFL
VDNSSVDFQKFPDKEILRMAGPLTADFI VKIRNSGSADSTVQFIFYQPIIHRWRETDFFP
ATCGGGYQLTSAECYDLRSNRVVADQYCHYYPENIKPKPKLQECNLDPASDGYKQIMPYD
LYHPLPRWEATPWTACSSSCGGGIQSRAVSCVEEDIQGHVTSVEEWKCMYTPKMPIAQPCNI
FDCPKWLAQEWSPCTVTGQGLRYRVVLCIDHRGMHTGGCSPKTKPHIKEECIVPTPCYKPK
EKL PVEAKLPWFKQAQAELEEGA AVSEEPS

Important features:

Signal peptide:

amino acids 1-25

N-glycosylation site.

amino acids 251-254

Thrombospondin 1

amino acids 385-399

von Willebrand factor type C domain proteins

amino acids 385-399, 445-459 and 42-56

FIGURE 121

CGGACGCGTGGGCGGCGGCTGCGGAACTCCCGTGGAGGGGCCGGTGGGCCCTCGGGCCTGAC
AGATGGCAGTGGCCACTGCGGCGGCAGTACTGGCCGCTCTGGGCGGGGCGCTGTGGCTGGCG
GCCCCCGGTTTCGTGGGGCCCAGGGTCCAGCGGCTGCGCAGAGGCGGGGACCCCGGCCTCAT
GCACGGGAAGACTGTGCTGATCACCGGGGCGAACAGCGGCCTGGGCCGCGCCACGGCCGCCG
AGCTACTGCGCCTGGGAGCGCGGGTGATCATGGGCTGCCGGGACCGCGCGCGCGCCGAGGAG
GCGGCGGGTCAGCTCCGCCGCGAGCTCCGCCAGGCCGCGGAGTGCGGCCAGAGCCTGGCGT
CAGCGGGGTGGGCGAGCTCATAGTCCGGGAGCTGGACCTCGCCTCGCTGCGCTCGGTGCGCG
CCTTCTGCCAGGAAATGCTCCAGGAAGAGCCTAGGCTGGATGTCTTGATCAATAACGCAGGG
ATCTTCCAGTGCCCTTACATGAAGACTGAAGATGGGTTTGAGATGCAGTTCGGAGTGAACCA
TCTGGGGCACTTTCTACTCACCAATCTTCTCCTTGGA CTCTCAAAGTTTCAGCTCCCAGCA
GGATTGTGGTAGTTTTCTTCCAACTTTATAAATACGGAGACATCAATTTTGATGACTTGAAC
AGTGAACAAAGCTATAATAAAAGCTTTTGTATAGCCGGAGCAAAGTGGCTAACATTCTTTT
TACCAGGGAAGTACCCGCCGCTTAGAAGGCACAAATGTCACCGTCAATGTGTGTCATCCTG
GTATTGTACGGACAAATCTGGGGAGGCACATACACATTCCACTGTTGGTCAAACCACTCTTC
AATTTGGTGTGTCATGGGCTTTTTTCAAAGTCCAGTAGAAGGTGCCCAGACTTCCATTTATTT
GGCCTCTTCACCTGAGGTAGAAGGAGTGTGAGGAAGATACTTTGGGGATTGTAAAGAGGAAG
AACTGTTGCCCAAAGCTATGGATGAATCTGTTGCAAGAAAGTCTGGGATATCAGTGAAGTG
ATGGTTGGCCTGCTAAAATAGGAACAAGGAGTAAAAGAGCTGTTTATAAAAGTGCATATCAG
TTATATCTGTGATCAGGAATGGTGTGGATTGAGAACTTGTTACTTGAAGAAAAGAATTTTG
ATATTGGAATAGCCTGCTAAGAGGTACATGTGGGTATTTTGGAGTTACTGAAAAATTATTTT
TGGGATAAGAGAATTTTCAGCAAAGATGTTTTAAATATATATAGTAAGTATAATGAATAATA
GTACAATGAAAAATACAATTATATTGTAAAATTATAACTGGGCAAGCATGGATGACATATTA
ATATTTGTCAGAAATTAAGTGACTCAAAGTGCTATCGAGAGGTTTTTCAAGTATCTTTGAGTT
TCATGGCCAAAGTGTTAACTAGTTTTTACTACAATGTTTGGTGTGTTGTGTGGAAATTATCTGC
CTGGTGTGTGCACACAAGTCTTACTTGGAATAAATTTACTGGTAC

FIGURE 122

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA58747

<subunit 1 of 1, 336 aa, 1 stop

<MW: 36865, pI: 9.15, NX(S/T): 2

MAVATAAAVLAAALGGALWLAARRFVGPRVQRLRRGGDPGLMHGKTVLITGANSGLGRATAAE
LLRLGARVIMGCRDRARAEAAAGQLRREL RQAAECGPEPGVSGVGELIVRELDLASLRSVRA
FCQEMLQEEPRLDVLINNAGIFQCPYMKTEDGFEMQFGVNHLGHFLLTNLLLGLLKSSAPSR
IVVVSSKLYKYGDINFDDLNSEQSYNKSFCYSRSLANILFTRELARRLEGTNVTNVNLHPG
IVRTNLGRHIHIPLLVKPLFNLVSWAFFKTPVEGAQTSIYLASSPEVEGVSGRYFGDCKEEE
LLPKAMDESVARKLWDISEVMVGLLK

Important features:

Signal peptide:

amino acids 1-21

Short-chain alcohol dehydrogenase family protein

amino acids 134-144, 44-56 and 239-248

N-glycosylation site.

amino acids 212-215 and 239-242

FIGURE 123

GGGGATTGTAAAGAGGAAGNACTGTGCCCAAAGNTATGGATGAATCTGTTGCAAGAAAATTN
TGGGATATCAGTGAAGTGATGGTTNGCCTGCTAAAATAGGAACAAGGAGTAAAAGAGCTGTT
TATAAACTGCATATCAGTTATATCTGTGATCAGGAATGGTGTGGATTGAGAACTTGTTACT
TGAAGAAAAAGAATTTTGATATTGGAATAGCCTGNTAAGAGGNACATGTGGGTATTTTGGAG
TTACTGAAAAATTATTTTGGGATAAGAGAATTCAGCAAAGATGTTTTAAATATATATAGT
AAGTATAATGAATAATAAGTACAATGAAAAATACAATTATATTGTAAAATTATAACTGGGCA
AGCATGGATGACATATTAATATTTGTCAGAATTAAGTGACTCAAAGTGCTATCGAGAGGTTT
TTCAAGTATCTTTGAGTTTCATGGCCAAAGTGTTAACTAGTTTTACTACAATGTTTGGTGTT
TGTGTGGAAATTATCTGCCTGGCTT

FIGURE 124

GAGAGGACGAGGTGCCGCTGCCTGGAGAATCCTCCGCTGCCGTCGGCTCCCGGAGCCCAGCC
CTTTCCTAACCCAACCCAACCTAGCCCAGTCCCAGCCGCCAGCGCCTGTCCCTGTCACGGAC
CCCAGCGTTACCATGCATCCTGCCGTCTTCCTATCCTTACCCGACCTCAGATGCTCCCTTCT
GCTCCTGGTAACTTGGGTTTTTACTCCTGTAACAACTGAAATAACAAGTCTTGCTACAGAGA
ATATAGATGAAATTTTAAACAATGCTGATGTTGCTTTAGTAAATTTTTATGCTGACTGGTGT
CGTTTCAGTCAGATGTTGCATCCAATTTTTGAGGAAGCTTCCGATGTCATTAAGGAAGAATT
TCCAAATGAAAATCAAGTAGTGTTTGCCAGAGTTGATTGTGATCAGCACTCTGACATAGCCC
AGAGATACAGGATAAGCAAATACCCAACCTCAAATTGTTTCGTAATGGGATGATGATGAAG
AGAGAATACAGGGGTGAGCGATCAGTGAAAGCATTGGCAGATTACATCAGGCAACAAAAAAG
TGACCCCATTCAGAAATTCGGGACTTAGCAGAAATCACCACTCTTGATCGCAGCAAAAGAA
ATATCATTGGATATTTTTGAGCAAAAGGACTCGGACAACTATAGAGTTTTTGAACGAGTAGCG
AATATTTTGCATGATGACTGTGCCTTTCTTTCTGCATTTGGGGATGTTTCAAACCGGAAAG
ATATAGTGGCGACAACATAATCTACAAACCACCAGGGCATTTCTGCTCCGGATATGGTGTACT
TGGGAGCTATGACAAATTTTGATGTGACTTACAATTGGATTCAAGATAAATGTGTTCTCTTT
GTCCGAGAAATAACATTTGAAAATGGAGAGGAATTGACAGAAGAAGGACTGCCTTTTCTCAT
ACTCTTTCACATGAAAGAAGATACAGAAAGTTTAGAAATATTCCAGAATGAAGTAGCTCGGC
AATTAATAAGTGAAAAAGGTACAATAAACTTTTTACATGCCGATTGTGACAAATTTAGACAT
CCTCTTCTGCACATACAGAAAACCTCCAGCAGATTGTCCTGTAATCGCTATTGACAGCTTTAG
GCATATGTATGTGTTTGGAGACTTCAAAGATGTATTAATTCTTGGAAGCACTCAAGCAATTCG
TATTTGACTTACATTCTGGAAAACCTGCACAGAGAATTCCATCATGGACCTGACCCAACTGAT
ACAGCCCCAGGAGAGCAAGCCCAAGATGTAGCAAGCAGTCCACCTGAGAGCTCCTTCCAGAA
ACTAGCACCCAGTGAATATAGGTATACTCTATTGAGGGATCGAGATGAGCTTTTAAAAAACTTG
AAAAACAGTTTGTAAGCCTTTCAACAGCAGCATCAACCTACGTGGTGGAAATAGTAAACCTA
TATTTTCATAATTCTATGTGTATTTTTATTTTGAATAAACAGAAAGAAATTTAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 125

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA57689

<subunit 1 of 1, 406 aa, 1 stop

<MW: 46927, pI: 5.21, NX(S/T): 0

MHPAVFLSLPDLRCSLLLLVTWVFTPVTTEITSLATENIDEILNNADVALVNFYADWCRFSQ
MLHPIFEEASDVIKEEFPNENQVVFARVDCDQHSDIAQRYRISKYPTLKLFRNGMMMKREYR
GQRSVKALADYIRQQKSDPIQEIRDLAEITTLDRSKRNIIGYFEQKDSDNRYRVFERVANILH
DDCAFLSAFGDVSKPERYSGDNIIYKPPGHSAPDMVYLGAMTNFDVTYNWIQDKCVPLVREI
TFENGEEELTEEGLPFLILFHMKEDESLEIFQNEVARQLISEKGTINFLHADCDKFRHPLLH
IQKTPADCPVIAIDSFRHMYVFGDFKDVLIPGKLKQFVFDLHSGKLHREFHHGPDPTDTAPG
EQAQDVASSPPESSFQKLAPSEYRYTLLRDRDEL

Important features:

Signal peptide:

amino acids 1-29

Endoplasmic reticulum targeting sequence.

amino acids 403-406

Tyrosine kinase phosphorylation site.

amino acids 203-211

Thioredoxin family proteins

amino acids 50-66

FIGURE 126

ATTAAGGAAGAATTTCCAAATGAAAATCAAGTAGTNTTTGCCAGAGTNGATTGTGATCAGCA
CTCTGACATAGCCCAGAGATACAGGATAAGCAAATACCCAACCCTCAAATTGTTTCGTAATG
GGATGATGATGAAGAGAGAATACAGGGGTCAGCGATCAGTGAAAGCATTGGCAGATTA

FIGURE 127

AGAGGCCTCTCTGGAAGTTGTCCCGGGTGTTGCGCCGCGNGGAGCCCCGGGTCGAGAGGACNAGG
TGCCGCTGCCTGGAGAATCCTCCGCTGCCGTGCGCTCCCGGAGCCCAGCCCTTTCCTAACCC
AACCCAACCTAGCCCNGTCCCAGCCGCCAGCGCCTGTCCCTGTCNCGGANCCCAGCGTNACC
ATGCATCCTGCCGTCTTCCTATCCTTACCCGACCTCAGATGCTCCCTTCTGCTCCTGGTAAC
TTGGGTTTTTACTCCTGTAACTGAAATAACNNGTCTTGATACNNAGAATATAGATGAAA
TTTTAAACNATGCTGATGTGGCTTTAGTCAATTTTTATGCTGACTGGTGTCGTTTCAGTCAG
ATGTGGCATCCAATTTTTGAGGANGCTTCCGATGTCATTAAGGAAGAATTTCCAAATGAAAA
TCAAGTAGTGTTTGCCAGAGTTGATTGTGATCAGCACTCTGACATAGCCCAGAGATACAGGA
TAAGCAAATACCCAACCCTCAAATTGTTTCGTAATGGGATGATGATGAAGAGAGAATACAGG
GGTCAGCGATCAGTGAAAGCATTGGCAGATTACATCAGGC

FIGURE 128

GCCCCACGCGTCCGATGGCGTTCACGTTGCGGGCCTTCTGCTACATGCTGGCGCTGCTGCTCA
CTGCCGCGCTCATCTTCTTCGCCATTTGGCACATTATAGCATTGATGAGCTGAAGACTGAT
TACAAGAATCCTATAGACCAGTGTAATACCCTGAATCCCCTTGTAAGTCCCAGAGTACCTCAT
CCACGCTTTCTTCTGTGTGTCATGTTTCTTTGTGCAGCAGAGTGGCTTACACTGGGTCTCAATA
TGCCCCCTCTTGGCATATCATATTTGGAGGTATATGAGTAGACCAGTGATGAGTGGCCCAGGA
CTCTATGACCCTACAACCATCATGAATGCAGATATTCTAGCATATTGTCAGAAGGAAGGATG
GTGCAAATTAGCTTTTTATCTTCTAGCATTTTTTTTACTACCTATATGGCATGATCTATGTTT
TGGTGAGCTCTTAGAACAACACACAGAAGAATTGGTCCAGTTAAGTGCATGCAAAAAGCCAC
CAAATGAAGGGATTCTATCCAGCAAGATCCTGTCCAAGAGTAGCCTGTGGAATCTGATCAGT
TACTTTAAAAAATGACTCCTTATTTTTTTAAATGTTTCCACATTTTTTGCTTGTGGAAAGACTG
TTTTTCATATGTTATACTCAGATAAAGATTTTTAAATGGTATTACGTATAAATTAATATAAAAT
GATTACCTCTGGTGTGACAGGTTTGAAGTTGCACTTCTTAAGGAACAGCCATAATCCTCTG
AATGATGCATTAATTACTGACTGTCTAGTACATTGGAAGCTTTTGTTTATAGGAAGTTGTA
GGGCTCATTTTTGGTTTTATTGAAACAGTATCTAATTATAAATTAGCTGTAGATATCAGGTGC
TTCTGATGAAGTGAAAATGTATATCTGACTAGTGGGAACTTCATGGGTTTCCTCATCTGTC
ATGTCGATGATTATATATGGATACATTTACAAAAATAAAAAGCGGGAATTTTCCCTTCGCTT
GAATATTATCCCTGTATATTGCATGAATGAGAGATTTCCCATATTTCCATCAGAGTAATAAA
TATACTTGCTTTAATTCTTAAGCATAAGTAAACATGATATAAAAATATATGCTGAATTACTT
GTGAAGAATGCATTTAAAGCTATTTTTAAATGTGTTTTTATTTGTAAGACATTACTTATTAAG
AAATTGGTTATTATGCTTACTGTTCTAATCTGGTGGTAAAGGTATTCTTAAGAATTTGCAGG
TACTACAGATTTTCAAACTGAATGAGAGAAAATTGTATAACCATCCTGCTGTTTCTTTAGT
GCAATACAATAAACTCTGAAATTAAGACTC

FIGURE 129

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA23330

<subunit 1 of 1, 144 aa, 1 stop

<MW: 16699, pI: 5.60, NX(S/T): 0

MAFTFAAFCYMLALLLTAAALIFFAIWHIIAFDELKTDYKNPIDQCNTLNPLVLPEYLIHAFF
CVMFLCAAEWLTLGLNMPLLAYHIWRYMSRPVMSGPGLYDPTTIMNADILAYCQKEGWCKLA
FYLLAFFYYLYGMIYVLVSS

Important features:

Signal peptide:

amino acids 1-20

Type II transmembrane domain:

amino acids 11-31

Other transmembrane domain:

amino acids 57-77 and 123-143

FIGURE 130

ATTATAGCATTTGATGAGCTGAAGACTGATTACAAGATCCTATAGACCAGTGTAATACCCTG
AATCCCCTTGTA TACTCCCAGAGTACCTCATCCACGCTTTCTTCTGTGTCATGTTTCTTTGTGC
AGCAGAGTGGCTTACACTGGGTCTCAATATGCCCCTCTTGGCATATCATATTTGGAGGTATA
TGAGTAGACCAGTGATGAGTGGCCCAGGACTCTATGACCCTACAACCATCATGAATGCAGAT
ATTCTAGCATATTGTCAGAAGGAAGGATGGTGCAAATTAGCTTTTTATCTTCTAGCATTTTT
TTACTACCTATATGGCATGATCTATGTTTTGGTGAGCTCTTAGAACAACACACAGAAGAATT
GGTCCAGTTAAGTG CATGCAAAAAGCCACCAAATGAAGGGATTCTATCCAGCAAGATCCTGT
CCAAGAGTAGCCTGTGGAATCTGATCAGTTACTTTAAAAAATG

FIGURE 131

CGGACGCGTGGGGGAAACCCTTCCGAGAAAACAGCAACAAGCTGAGCTGCTGTGACAGAGGG
GAACAAGATGGCGGCGCCGAAGGGGAGCCTCTGGGTGAGGACCCAACTGGGGCTCCCGCCGC
TGCTGCTGCTGACCATGGCCTTGGCCGGAGGTTGCGGGACCGCTTCGGCTGAAGCATTGAC
TCGGTCTTGGGTGATACGGCGTCTTGCCACCGGGCCTGTCAGTTGACCTACCCCTTGACAC
CTACCCTAAGGAAGAGGAGTTGTACGCATGTCAGAGAGGTTGCAGGCTGTTTTCAATTTGTC
AGTTTGTGGATGATGGAATTGACTTAAATCGAACTAAATTGGAATGTGAATCTGCATGTACA
GAAGCATATTCCCAATCTGATGAGCAATATGCTTGCCATCTTGGTTGCCAGAATCAGCTGCC
ATTTCGCTGAACTGAGACAAGAACAACCTTATGTCCCTGATGCCAAAAATGCACCTACTCTTTC
CTCTAACTCTGGTGAGGTCATTCTGGAGTGACATGATGGACTCCGCACAGAGCTTCATAACC
TCTTCATGGACTTTTTTATCTTCAAGCCGATGACGGAAAAATAGTTATATTCCAGTCTAAGCC
AGAAATCCAGTACGCACCACATTTGGAGCAGGAGCCTACAAATTTGAGAGAATCATCTCTAA
GCAAAATGTCCTATCTGCAAATGAGAAATTCACAAGCGCACAGGAATTTTCTTGAAGATGGA
GAAAGTGATGGCTTTTTTAAGATGCCTCTCTCTTAACCTCTGGGTGGATTTTAACTACAACCTCT
TGTCCTCTCGGTGATGGTATTGCTTTGGATTTGTTGTGCAACTGTTGCTACAGCTGTGGAGC
AGTATGTTCCCTCTGAGAAGCTGAGTATCTATGCTGACTTGGAGTTTATGAATGAACAAAAG
CTAAACAGATATCCAGCTTCTTCTCTTGTGGTTGTTAGATCTAAACTGAAGATCATGAAGA
AGCAGGGCCTCTACCTACAAAAGTGAATCTTGCTCATCTGAAATTTAAGCATTTTTCTTTT
AAAAGACAAGTGTAATAGACATCTAAATTCACCTCCTCATAGAGCTTTTAAAATGGTTTCA
TTGGATATAGGCCTTAAGAAATCACTATAAAATGCAAATAAAGTTACTCAAATCTGTG

FIGURE 132

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA26847

<subunit 1 of 1, 323 aa, 1 stop

<MW: 36223, pI: 5.06, NX(S/T): 1

MAAPKGSLWVRTQLGLPPLLLLTMALAGGSGTASAEAFDSVLGDTASCHRAQLTYPLHTYP
KEEELYACQRGCRLEFSICQFVDDGIDLNRKLECESACTEAYSQSDEQYACHLGCQNQLPFA
ELRQEQQLMSLMPKMHLFPLTLVRSFWSDMMDSAQSFITSSWTFYQLADDGKIVIFQSKPEI
QYAPHLEQEPTNLRESSLSKMSYQLMRNSQAHRNFLEDGESDGFLRCLSLNSGWILTTTLVL
SVMVLLWICCATVATAVEQYVPSEKLSIYGDLEFMNEQKLNRYPASSLVVVRSKTEDHEEAG
PLPTKVNLAHSEI

Important features:

Signal peptide:

amino acids 1-31

Transmembrane domain:

amino acids 241-260

N-glycosylation site.

amino acids 90-93

FIGURE 133

TTGGGTGATACGGCGTCTTGCCACCGGGCCTGTCAGTTGACCTACCCCTTGACACCTACCC
TAAGGAAGAGGAGTTGTACGCATGTCAGAGAGGTTGCAGGCTGTTTTCAATTGTCAGTTTG
TGGATGATGGAATTGACTTAAATCGAACTAAATTGGAATGTGAATCTGCATGTACAGAAGCA
TATTCCCAATCTGATGAGCAATATGCTTGCCATCTTGGTTGCCAGAATCAGCTGCCATTTCGC
TGAAGTGAAGACAAGAACAACCTTATGTCCCTGATGCCAAAAATGCACCTACTCTTTCCTCTAA
CTCTGGTGAGGTCATTCTGGAGTGACATGATGGACTCCGC

FIGURE 134

CACACTGGCCGGATCTTTTAGAGTCCTTTGACCTTGACCAAGGGTCNGGAAAACAGCAACAA
GCTGAGCTGCTGTGACAGAGGGAACAAGATGGCGGCGCCGAAGGGAGCCTTTGGGTGAGGAC
CCAAC TGGGGCTCCCGCCGCTGCTGCTGCTGACCATGGCCTTGGCCGGAGGTTGCGGGACCG
CTTCGGCTGAAGCATTTGACTCGGTCTTGGGTGATACGGCGTCTTGCCACCGGGCCTGTCAG
TTGACCTACCCCTTGACACCTACCCTAAGGAAGAGGAGTTGTACGCATGTCAGAGAGGTTG
CAGGCTGTTTTCAATTTGTCAGTTTGTGGATGATGGAATTGACTTAAATCGAACTAAATTGG
AATGTGAATCTGCATGTACAGAAGCATATTCCCAATCTGATGAGCAATATGCTTGCCATCTT
GGTTGCCAGAATCAGCTGCCATTGCTGAACTGAGACAAGAACAACCTTATGTCCCTGATGCC
AAAAATGCACCTACTCTTTCCTCTAACTCTGGTGAGGTCATTCTGGAGTGACATGATGGACT
CCGC

FIGURE 135

CGGAGGTGGCGATCGCTGAGAGGCAGGAGGGCCGAGGCGGGCCTGGGAGGCGGCCCGGAGGT
GGGGCGCCGCTGGGGCCGGCCCGCACGGGCTTCATCTGAGGGCGCACGGCCCGCGACCGAGC
GTGCGGACTGGCCTCCCAAGCGTGGGGCGACAAGCTGCCGGAGCTGCAATGGGCCGCGGCTG
GGGATTCTTGTTTGGCCTCCTGGGCGCCGTGTGGCTGCTCAGCTCGGGCCACGGAGAGGAGC
AGCCCCCGGAGACAGCGGCACAGAGGTGCTTCTGCCAGGTTAGTGGTTACTTGGATGATTGT
ACCTGTGATGTTGAAACCATTGATAGATTTAATAACTACAGGCTTTTCCCAAGACTACAAAA
ACTTCTTGAAAGTGACTACTTTAGGTATTACAAGGTAAACCTGAAGAGGCCGTGTCTTTCT
GGAATGACATCAGCCAGTGTGGAAGAAGGGACTGTGCTGTCAAACCATGTCAATCTGATGAA
GTTCTGTATGGAATTAAATCTGCGAGCTACAAGTATTCTGAAGAAGCCAATAATCTCATTGA
AGAATGTGAACAAGCTGAACGACTTGGAGCAGTGGATGAATCTCTGAGTGAGGAAACACAGA
AGGCTGTTCTTCAGTGGACCAAGCATGATGATTCTTCAGATAACTTCTGTGAAGCTGATGAC
ATTCAGTCCCCCTGAAGCTGAATATGTAGATTTGCTTCTTAATCCTGAGCGCTACACTGGTTA
CAAGGGACCAGATGCTTGGAATAATGGAATGTCATCTACGAAGAAAACTGTTTTAAGCCAC
AGACAATTAAAAGACCTTTAAATCCTTTGGCTTCTGGTCAAGGGACAAGTGAAGAGAACT
TTTTACAGTTGGCTAGAAGGTCTCTGTGTAGAAAAAAGAGCATTCTACAGACTTATATCTGG
CCTACATGCAAGCATTAAATGTGCATTTGAGTGCAAGATATCTTTTACAAGAGACCTGGTTAG
AAAAGAAATGGGGACACAACATTACAGAATTTCAACAGCGATTTGATGGAATTTTGACTGAA
GGAGAAGTCCAAGAAGGCTTAAGAACTTGTATTTTCTCTACTTAATAGAACTAAGGGCTTT
ATCCAAAGTGTTACCATTTCTTCGAGCGCCAGATTTTCAACTCTTTACTGGAAATAAAATTC
AGGATGAGGAAAAACAAAATGTTACTTCTGGAAATACTTCATGAAATCAAGTCATTTCTTTTG
CATTTTGATGAGAATTCATTTTTTGTCTGGGGATAAAAAAGAAGCACACAACTAAAGGAGGA
CTTTTCGACTGCATTTTAGAAATATTTCAAGAATTATGGATTGTGTTGGTTGTTTTAAATGTC
GTCTGTGGGGAAAGCTTCAGACTCAGGGTTTGGGCACTGCTCTGAAGATCTTATTTTCTGAG
AAATTGATAGCAAATATGCCAGAAAGTGGACCTAGTTATGAATCCATCTAACCAGACAAGA
AATAGTATCATTATTCAACGCATTTGGAAGAATTTCTACAAGTGTGAAAGAATTAGAAAAC
TCAGGAACTTGTTACAGAATATTCATTAAAGAAAACAAGCTGATATGTGCCTGTTTCTGGAC
AATGGAGGCGAAAGAGTGGAATTTCAATCAAAGGCATAATAGCAATGACAGTCTTAAGCCAA
ACATTTTATATAAAGTTGCTTTTGTAAAGGAGAATTATATTGTTTTAAGTAAACACATTTTT
AAAAATTGTGTTAAGTCTATGTATAATACTACTGTGAGTAAAGTAATACTTTAATAATGTG
GTACAAATTTTAAAGTTTAAATATTGAATAAAAGGAGGATTATCAAATTAATAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAA

[illegible]

<subunit 1 of 1, 468 aa, 1 stop

MGRGWGFLFGLLGAVWLLSSGHGEEQPPETAARCFQCQVSGYLDDCTCDVETIDRFNNYRLF
PRLQKLLESDYFRYYKVNLRPCPFWNDISQCGRRDCAVKPCQSDEVPDGIKSASYKYSEEA
NNLIEECEQAERLGAVDESLSSEETQKAVLQWTKHDDSSDNFCEADDIQSPEAEYVDLLLNP
RYTGYKGPDAWKIWNVIYEENCFKPQTIKRPLNPLASGQGTSEENTFYSWLEGLCVEKRAF
RLISGLHASINVHLSARYLLQETWLEKKWGHNITEFQQRFDGILTEGEGPRRLKNLYFLYLI
ELRALSKVLPPFFERPDLFTGNKIQDEENKMLLLEILHEIKSFPLHFDENSFFAGDKKEAH
KLKEDFRLHFRNISRIMDCVGCFCRLWGKLQTQGLGTALKILFSEKLIANMPESGPSYEFH
LTRQEI VSLFNAFGRISTSVKELENERNLLONIH

Signal peptide:

N-glycosylation site.

Amidation site.

Glycosaminoglycan attachment site.

amino acids 20-23 and 223-226

Aminotransferases class-V pyridoxal-phosphate

amino acids 216-222

Interleukin-7 proteins

amino acids 338-343

FIGURE 137

GCTGGAAATATGGATGTCATCTACGAGAACTGTTTTAAGCCACAGACAATTAAAAGACCTT
TAAATCCTTTGGCTTCTGGTCAAGGGACAAGTGAAGAGNACACTTTTTACAGTTGGCTAGAA
GGTCTCTGTGTAGAAAAAGAGCATTCTACAGACTTATATCTGGCCTACATGCAAGCATTAA
TGTGCATTTGAGTGCAAGATATCTTTTACAAGAGACCTGGTTAGAAAAGAAATGGGGACACA
ACATTACAGAATTTNAACAGCGATTTGATGGAATTTTGACTGAAGGAGAAGGTCCAAGAAGG
CTTAAGAACTTGTATTTTCTCTACTTAATAGAACTAAGGGCTTTATCCAAAGTGTTACCATT
CTTNGAGCGCCCAGATTTTCAACTNTTTACTGGAAATAAAATTCAGGATGAGGNAAACAAAA
TGTTACTTTTGGAATACTTCATGAAATCAAGTCATTTCTTTGCATTTTGATGAGAATTCA
TTTTTTTGCTG

FIGURE 138

CGGACGCGTGGGCGGACGCGTGGGCGGACGCGTGGGTTGGGAGGGGGCAGGATGGGAGGGAA
AGTGAAGAAAACAGAAAAGGAGAGGGACAGAGGCCAGAGGACTTCTCATACTGGACAGAAAC
CGATCAGGCATGGAACTCCCCCTTCGTCACTCACCTGTTCTTGCCCCCTGGTGTTCTTGACAGG
TCTCTGCTCCCCCTTTAACCTGGATGAACATCACCCACGCCTATTCCCAGGGGCCACCAGAAG
CTGAATTTGGATACAGTGTCTTACAACATGTTGGGGGTGGACAGCGATGGATGCTGGTGGGC
GCCCCCTGGGATGGGCCTTCAGGCGACCGGAGGGGGGACGTTTATCGCTGCCCTGTAGGGGG
GGCCACAAATGCCCCATGTGCCAAGGGCCACTTAGGTGACTACCAACTGGGAAATTCATCTC
ATCCTGCTGTGAATATGCACCTGGGGATGTCTCTGTTAGAGACAGATGGTGATGGGGGATTC
ATGGTGAGCTAAGGAGAGGGTGGTGGCAGTGTCTCTGAAGGTCCATAAAAGAAAAAGAGAA
GTGTGGTAAGGGAAAATGGTCTGTGTGGAGGGGTCAAGGAGTTAAAAACCCTAGAAAGCAAA
AGGTAGGTAATGTCAGGGAGTAGTCTTCATGCCTCCTTCAACTGGGAGCATGTTCTGAGGGT
GCCCTCCCAAGCCTGGGAGTAACTATTTCCCCCATCCCCAGGCCTGTGCCCTCTCTGGTCT
CGTGCTTGTGGCAGCTCTGTCTTCAGTTCCTGGGATATGTGCCCGTGTGGATGCTTCATTCCA
GCCTCAGGGAAGCCTGGCACCCACTGCCAACGTGAGCCAGAGGAAGGCTGAGTACTTGTTT
CCCAGAAGGAGATACTGGGTGGGAAAAAGATGGGGCAAAGCGGTATGATGCCTGGCAAAGGG
CCTGCATGGCTATCCTCATTGCTACCTAATGTGCTTGCAAAGCTCCATGTTTCCTAACAGA
TTCAGACTCCTGGCCAGGTGTGGTGGCCACACCTGTAATTCTAGCACTTTGGGAGGCCAAG
GTGGGCAGATCACTTGAGGTCAGGAGTTCAAGACCAGCCTGGCCAACATGGTGAAACTCCAT
CTCTACTAAAAAAAAAAAAATACAAAAATTAGCTGGGTGCGCTAGTGCATGCCTGTAATCTC
ATCTACTCGGGAGGCTAAGACAGGAGACTCTCACTTCAACCCAGGAGGTGGAGGTTGCGGTG
AGCCAAGATTGTGCCTCTGCACTCTAGCGTGGGTGACAGAGTAAGCGAGACTCCATCTCAA
AATAATAATAATAATAATTCAGACTCCTTATCAGGAGTCCATGATCTGGCCTGGCACAGTAA
CTCATGCCTGTAATCCCAACATTTTGGGAGGCCAACGCAGGAGGATTGCTTGAGGTCTGGAG
GTTTGAGACCAGCCTGGGCAACATAGAAAGACCCCATCTCTAAATAAATGTTTTAAAAAT

FIGURE 139

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA57039

><subunit 1 of 1, 124 aa, 1 stop

><MW: 13352, pI: 5.99, NX(S/T): 1

MELPFVTHLFLPLVFLTGLCSPFNLDEHHPRLFPGPPEAEFGYSVLQHVGGGQRWMLVGAPW
DGPSGDRRGDVYRCPVGGAHNAPCAKGHLGDYQLGNSSHPAVNMHLGMSLLETGDGGFMVS

Important features:

Signal peptide:

amino acids 1-22

Cell attachment sequence.

amino acids 70-73

N-glycosylation site.

amino acids 98-101

Integrins alpha chain proteins

amino acids 67-81

FIGURE 140

CACAGTTCCCCACCATCACTCNTCCCATTCCTTCCAACTTTATTTTTAGCTTGCCATTGGGA
GGGGGCAGGATGGGAGGGAAAGTGAAGAAAACAGAAAAGGAGAGGGACAGAGGCCAGAGGAC
TTCTCATACTGGACAGAAACCGATCAGGCATGGAACTCCCCTTCGTCACTCACCTGTTCTTG
CCCCTGGTGTTTCCTGACAGGTCTCTGCTCCCCCTTTAACCTGGATGAACATCACCCACGCCT
ATTCCCAGGGCCACCAGAAGCTGAATTTGGATACAGTGTCTTACAACATGTTGGGGGTGGAC
AGCGATGGATGCTGGTGGGCGCCCCCTGGGATGGGCCTTCAGGCGACCGGAGGGGGGACGTT
TATCGCTGCCCTGTAGGGGGGGCCCAATGCCCCATGTGCCAAGGGCCACTTAGGTGACTA
CCAACTGGGAAATTCATCTCATCCTGCTGTGAATATGCACCTGGGGATGTCTCTGTTAGAGA
CAGATGGTGATGG

FIGURE 141

AAAGTTACATTTTCTCTGGAACCTCTCCTAGGCCACTCCCTGCTGATGCAACATCTGGGTTTG
GGCAGAAAGGAGGGTGCTTCGGAGCCCGCCCTTTCTGAGCTTCCTGGGCCGGCTCTAGAACA
ATTGAGGCTTCGCTGCGACTCAGACCTCAGCTCCAACATATGCATTCTGAAGAAAGATGGCT
GAGATGGACAGAATGCTTTATTTTGGAAAGAAACAATGTTCTAGGTCAAACCTGAGTCTACCA
AATGCAGACTTTCACAATGGTTCTAGAAGAAATCTGGACAAGTCTTTTCATGTGGTTTTTCT
ACGCATTGATTCCATGTTTGCTCACAGATGAAGTGGCCATTCTGCCTGCCCTCAGAACCTC
TCTGTACTCTCAACCAACATGAAGCATCTCTTGATGTGGAGCCCAGTGATCGCGCCTGGAGA
AACAGTGTACTATTCTGTGCAATACCAGGGGGAGTACGAGAGCCTGTACACGAGCCACATCT
GGATCCCCCAGCAGCTGGTGCTCACTCACTGAAGGTCTGAGTGTGATGTCACTGATGACATC
ACGGCCACTGTGCCATACAACCTTCGTGTGAGGGCCACATTGGGCTCACAGACCTCAGCCTG
GAGCATCCTGAAGCATCCCTTTAATAGAACTCAACCATCCTTACCCGACCTGGGATGGAGA
TCACCAAAGATGGCTTCCACCTGGTTATTGAGCTGGAGGACCTGGGGCCCCAGTTTGAGTTC
CTTGTTGGCCTACTGGAGGAGGGAGCCTGGTGCCGAGGAACATGTCAAATGGTGAGGAGTGG
GGGTATTCCAGTGCACCTAGAAACCATGGAGCCAGGGGCTGCATACTGTGTGAAGGCCCAGA
CATTCGTGAAGGCCATTGGGAGGTACAGCGCCTTCAGCCAGACAGAATGTGTGGAGGTGCAA
GGAGAGGCCATTCCCCTGGTACTGGCCCTGTTTGCCTTTGTTGGCTTCATGCTGATCCTTGT
GGTCGTGCCACTGTTTCGTCTGGAATAATGGGCCGGCTGCTCCAGTACTCCTGTTGCCCGTGG
TGGTCTCCCGAGACACCTTGAAAATAACCAATTCACCCAGAAAGTTAATCAGCTGCAGAAGG
GAGGAGGTGGATGCCTGTGCCACGGCTGTGATGTCTCCTGAGGAACTCCTCAGGGCCTGGAT
CTCATAGGTTTTGCGGAAGGGCCCAGGTGAAGCCGAGAACCTGGTCTGCATGACATGGAAACC
ATGAGGGGACAAGTTGTGTTTCTGTTTTCCGCCACGGACAAGGGATGAGAGAAGTAGGAAGA
GCCTGTTGTCTACAAGTCTAGAAGCAACCATCAGAGGCAGGGTGGTTTTGTCTAACAGAACAC
TGAAGGAGGCTTAGGGGATGTGACCTCTAGACTGGGGGCTGCCACTTGCTGGCTGAGCAACC
CTGGGAAAAGTGACTTCATCCCTTCGGTCCCTAAGTTTTCTCATCTGTAATGGGGGAATTACC
TACACACCTGCTAAACACACACACAGAGTCTCTCTCTATATATACACACGTACACATAAA
TACACCCAGCACTTGCAAGGCTAGAGGGAACTGGTGACACTCTACAGTCTGACTGATTCAG
TGTTTCTGGAGAGCAGGACATAAATGTATGATGAGAATGATCAAGGACTCTACACACTGGGT
GGCTTGAGAGAGCCACTTTCCTCAGAATAATCCTTGAGAGAAAAGGAATCATGGGAGCAATGG
TGTTGAGTTCACCTTCAAGCCCAATGCCGGTGCAGAGGGGAATGGCTTAGCGAGCTCTACAGT
AGGTGACCTGGAGGAAGGTCACAGCCACACTGAAAATGGGATGTGCATGAACACGGAGGATC
CATGAACTACTGTAAAGTGTGACAGTGTGTGCACACTGCAGACAGCAGGTGAAATGTATGT
GTGCAATGCGACGAGAATGCAGAAGTCAGTAACATGTGCATGTTTGTGTGCTCCTTTTTTC
TGTGGTAAAGTACAGAATTGAGCAATAAAAAGGGCCACCCTGGCCAAAAGCGGTAAAAAA
AAAAAAAAAA

FIGURE 142

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA57033

<subunit 1 of 1, 311 aa, 1 stop

<MW: 35076, pI: 5.04, NX(S/T): 2

MQTFTMVLEEIWTSLFMWFFYALIPCLLTDEVAILPAPQNLSVLSTNMKHLIMWSPVIAPGE
TVYYSVEYQGEYESLYTSHIWIPSSWCSLTEGPECDVTDDITATVPYNLRVRATLGSQTS
SILKHPFNRNSTILTRPGMEITKDGFLVIELEDLGPQFEFLVAYWRREPGAEHVKMVRSG
GIPVHLETMEPGAAYCVKAQTFVKAIGRYSAFSQTECEVQGEAIPVLALFAFVGFMILILV
VVPLFVWKMGRLLQYSCCPVVVLPDTLKITNSPQKLISCRREEVDACATAVMSPEELLRAWIS

Important features:

Signal peptide:

amino acids 1-29

Transmembrane domain:

amino acids 230-255

N-glycosylation site.

amino acids 40-43 and 134-137

Tissue factor proteins.

amino acids 92-119

Integrins alpha chain proteins

amino acids 232-262

FIGURE 143

TCCTGCTGATGCACATCTGGGTTTGGCAAAAGGAGGTTGCTTCGAGCCGCCCTTTCTAGCTT
CCTGGCCGGCTCTAGAACAATTCAGGCTTCGCTGCGACTAGACCTCAGCTCCAACATATGCA
TTCTGAAGAAAGATGGCTGAGATGACAGAATGCTTTATTTTGGAAAGAAACAATGTTCTAGG
TCAAACCTGAGTCTACCAAATGCAGACTTTTCACAATGGTTCTAGAAGAAATCTGGACAAGTCT
TTTCATGTGGTTTTTCTACGCATTGATTCCATGTTTGCTCACAGATGAAGTGGCCATTCTGC
CTGCCCCCTCAGAACCTCTCTGTACTCTCAACCAACATGAAGCATCTCTTGATGTGGAGCCCA
GTGATCGCGCCTGGAGAAACAGTGTACTATTCTGTGCAATACCAGGGGGAGTACGAGAGCCT
GTACACGAGCCACATCTGGATCCCCAGCAGCTGGTGCTCACTCACTGAAGGTCTTGAGTGTG
ATGTCACTGATGACATCACGGCCACTGTGCCATACAACCTTTGTGTCAGGGCCACATTGGGC
TCACAGACCTCAGCCTGGAGCATCCTGAAGCATCCCTTTAATAGAAACTCAACCATCCTTAC
CCGACCTGGGATGGAGATCACCAAAGATGGCTTNCACCTGGTTATTGAGCTGGAGGACCTGG
GGCCCCAGTTTGAGTTCCTTGTGGCCTANTGGAGGAGGGGCGAACCCCTTGCGGCGCAAGGG
GTTNGCGAACCCCTTGCGGCCGCTGGGGTATCTCTCGAGAAAAGAGAGGCCCAATATGACCC
ACATACTCAATATGGACGAANTGCTATTGTCCACCTGTTTGAGTGGCGCTGGGTTGAT

FIGURE 144

CCCACGCGTCCGCCCACGCGTCCGAGGGACAAGAGAGAAGAGAGACTGAAACAGGGAGAAGA
GGCAGGAGAGGAGGAGGTGGGGAGAGCACGAAGCTGGAGGCCGACACTGAGGGAGGGCGGGA
GGAGGTGAAGAAGGAGAGAGGGGAGAAGAGGCAGGAGCTGGAAAGGAGAGAGGGAGGAGGAG
GAGGAGATGCGGGATGGAGACCTGGAGTTAGGTGGCTTGGGAGAGCTTAATGAAAAGAGAAC
GGAGAGGAGGTGTGGGTAGGAACCAAGAGGTAGCCCTGTGGGCAGCAGAAGGCTGAGAGGA
GTAGGAAGATCAGGAGCTAGAGGGAGACTGGAGGGTTCCGGGAAAAGAGCAGAGGAAAGAGG
AAAGACACAGAGAGACGGGAGAGAGAAGAAGAGTGGGTTTGAAGGGCGGATCTCAGTCCCTG
GCTGCTTTGGCATTGTTGGGAACTGGGACTCCCTGTGGGGAGGAGAGGAAAGCTGGAAGTCCT
GGAGGGACAGGGTCCCAGAAGGAGGGGACAGAGGAGCTGAGAGAGGGGGGACAGGGCGTTGGG
CAGGGGTCCCTCGGAGGCCTCCTGGGGATGGGGGCTGCAGCTCGTCTGAGCGCCCCCTCGAGC
GCTGGTACTCTGGGCTGCACTGGGGGCAGCAGCTCACATCGGACCAGCACCTGACCCCGAGG
ACTGGTGGAGCTACAAGGATAATCTCCAGGGAACTTCGTGCCAGGGCCTCCTTTCTGGGGC
CTGGTGAATGCAGCGTGGAGTCTGTGTGCTGTGGGGAAGCGGCAGAGCCCCGTGGATGTGGA
GCTGAAGAGGGTTCTTTATGACCCCTTTCTGCCCCCATTAAAGGCTCAGCACTGGAGGAGAGA
AGCTCCGGGGAACCTTGTACAACACCGGCCGACATGTCTCCTTCCTGCCTGCACCCCGACCT
GTGGTCAATGTGTCTGGAGGTCCCTCCTTTACAGCCACCGACTCAGTGAAGTGC GGCTGCT
GTTTGGAGCTCGCGACGGAGCCGGCTCGGAACATCAGATCAACCACCAGGGCTTCTCTGCTG
AGGTGCAGCTCATTCACTTCAACCAGGAACCTCTACGGGAATTTAGCGCTGCCTCCCGCGGC
CCCAATGGCCTGGCCATTCTCAGCCTCTTTGTCAACGTTGCCAGTACCTCTAACCATTCTC
CAGTCGCCTCCTTAACCGCGACACCATCACTCGCATCTCCTACAAGAATGATGCCTACTTTC
TTCAAGACCTGAGCCTGGAGCTCCTGTTCCCTGAATCCTTCGGCTTCATCACCTATCAGGGC
TCTCTCAGCACCCCGCCCTGCTCCGAGACTGTCACCTGGATCCTCATTGACCGGGCCCTCAA
TATCACCTCCCTTCAGATGCACTCCCTGAGACTCCTGAGCCAGAATCCTCCATCTCAGATCT
TCCAGAGCCTCAGCGGTAACAGCCGGCCCCCTGCAGCCCTTGGCCCACAGGGCACTGAGGGGC
AACAGGGACCCCCGGCACCCCGAGAGGCGCTGCCGAGGCCCACTACCGCCTGCATGTGGA
TGGTGTCCCCCATGGTCGCTGAGACTCCCTTTCGAGGATTGCACCCGCCCCGTCTAAGCCTC
CCCACAAGGCGAGGGGAGTTACCCCTAAAACAAAGCTATTAAAGGGACAGAATACTTA

FIGURE 145

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA34353

<subunit 1 of 1, 328 aa, 1 stop

<MW: 36238, pI: 9.90, NX(S/T): 3

MGAAARLSAPRALVLWAALGAAAHIGPAPDPEDWWSYKDNLQGNFVPGPPFWGLVNAAWSLC
AVGKRQSPVDVELKRVLYDPFLPPLRLSTGGEKLRGTLYNTGRHVSFLPAPRPVVNVSGGPL
LYSHRLSELRLLLFGARDGAGSEHQINHQGFSAEVQLIHFNQELYGNFSAASRGPNGLAILSL
FVNVASTSNPFLSRLLNRDTITRISYKNDAYFLQDLSLELLFPESFGFITYQGSLSSTPPCSE
TVTWILIDRALNITSLQMHSRLRLLSQNPPSQIFQSLSGNSRPLQPLAHRALRGNRDPRHPER
RCRGPNYRLHVDGVPHGR

Important features:

Signal peptide:

amino acids 1-23

Transmembrane domain:

amino acids 177-199

N-glycosylation site.

amino acids 118-121, 170-173 and 260-263

Eukaryotic-type carbonic anhydrases proteins

amino acids 222-270, 128-164 and 45-92

FIGURE 147

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA45417

<subunit 1 of 1, 500 aa, 1 stop

<MW: 56888, pI: 8.53, NX(S/T): 2

MKCTAREWLRVTTVLFMARAI PAMVVPNATLLEKLLLEKYMDEDEGEWWIAKQRGKRAITDNDM
QSILDLHNKLRSQVYPTASNMEYMTWDVELERSAESWAESCLWEHGPASLLPSIGQNLGAHW
GRYRPPTFHVQSWYDEVKDFSYPYEHECNPYCPFRCSGPVCTHYTQVWATSNRIGCAINLC
HNMNIWGQIWPKAVYLV CNYSPKGNWWGHAPYKHGRPCSACPPSFGGGCRENL CYKEGSDRY
YPPREEETNEIERQQSQVHDTHVRTRSDSSRNEVIS AQQMSQIVSCEVRLRDQCKGTT CNR
YEC PAGCLDSKAKVIGSVHYEMQSSICRAAIHYGIIDNDGGWVDITRQGRKH YFIKSNRNGI
QTIGKYQSANSFTVSKVTVQAVTCETTVEQLCPFH KPA SHCPRVYCPRNCMQANPHYARVIG
TRVYSDLSSICRAAVHAGVVRNHGGYVDVMPVDKRKTYIASFQNGIFSESLQNPPGGKA FRV
FAVV

Important features:

Signal peptide:

amino acids 1-20

Extracellular proteins SCP/Tpx-1/Ag5/PR-1/Sc7 protein

amino acids 165-186, 196-218, 134-146, 96-108 and 58-77

N-glycosylation site

amino acids 28-31

FIGURE 148

GCGGAGACAAGCGCAGAGCGCAGCGCACGGCCACAGACAGCCCTGGGCATCCACCGACGGCG
CAGCCGGAGCCAGCAGAGCCGGAAGGCGCGCCCCGGGCAGAGAAAGCCGAGCAGAGCTGGGT
GGCGTCTCCGGGCCCGCGCTCCGACGGGCCAGCGCCCTCCCCATGTCCCTGCTCCACGCCG
CGCCCCCTCCGGTCAGCATGAGGCTCCTGGCGGCCGCGCTGCTCCTGCTGCTGCTGGCGCTGT
ACACCGCGCGTGTGGACGGGTCCAAATGCAAGTGCTCCCGGAAGGGACCCAAGATCCGCTAC
AGCGACGTGAAGAAGCTGGAAATGAAGCCAAAGTACCCGCACTGCGAGGAGAAGATGGTTAT
CATCACCACCAAGAGCGTGTCCAGGTACCGAGGTCAGGAGCACTGCCTGCACCCCAAGCTGC
AGAGCACCAAGCGCTTCATCAAGTGGTACAACGCCTGGAACGAGAAGCGCAGGGTCTACGAA
GAATAGGGTGAAAAACCTCAGAAGGGAAAACTCCAAACCAGTTGGGAGACTTGTGCAAAGGA
CTTTGCAGATTAAAGCCTTTC
TTTCTCACAGGCATAAGACACAAATTATATATTGTTATGAAGCACTTTTTACCAACGGTCAG
TTTTTACATTTTATAGCTGCGTGCGAAAGGCTTCCAGATGGGAGACCCATCTCTCTTGTGCT
CCAGACTTCATCACAGGCTGCTTTTTATCAAAAAGGGGAAAACTCATGCCTTTCCTTTTAA
AAAATGCTTTTTTGTATTTGTCCATACGTCACTATACATCTGAGCTTTATAAGCGCCCGGA
GGAACAATGAGCTTGGTGGACACATTTCAATTGCAGTGTGCTCCATTCCTAGCTTGGGAAGC
TTCCGCTTAGAGGTCCTGGCGCCTCGGCACAGCTGCCACGGGCTCTCCTGGGCTTATGGCCG
GTCACAGCCTCAGTGTGACTCCACAGTGGCCCCGTGTAGCCGGGCAAGCAGGAGCAGGTCTCT
CTGCATCTGTTCTCTGAGGAACTCAAGTTTGGTTGCCAGAAAAATGTGCTTCATTCCCCCT
GGTTAATTTTTACACACCCTAGGAAACATTTCCAAGATCCTGTGATGGCGAGACAAATGATC
CTTAAAGAAGGTGTGGGGTCTTTCCCAACCTGAGGATTTCTGAAAGGTTACAGGTTCAATA
TTTAATGCTTCAGAAGCATGTGAGGTTCCCAACACTGTCAGCAAAAACCTTAGGAGAAAACT
TAAAAATATATGAATACATGCGCAATACACAGCTACAGACACACATTCTGTTGACAAGGGAA
AACCTTCAAAGCATGTTTCTTTCCCTCACCACAACAGAACATGCAGTACTAAAGCAATATAT
TTGTGATTCCCCATGTAATTCTTCAATGTTAAACAGTGCAGTCCTCTTTCGAAAGCTAAGAT
GACCATGCGCCCTTTCCTCTGTACATATACCCTTAAGAACGCCCCCTCCACACACTGCCCCC
CAGTATATGCCGCATTGTACTGCTGTGTTATATGCTATGTACATGTCAGAAACCATTAGCAT
TGCATGCAGGTTTCATATTCTTTCTAAGATGGAAAGTAATAAAATATATTTGAAATGTAAAA
AAAAAAAAAA

FIGURE 149

MSLLPRRAPPVSMRLLAAALLLLLLALYTARVDGSKCKCSRKGPKIRYSDVKKLEMKPKYPH
CEEKMVIITTKSVSRYRGQEHCLHPKLQSTKRFIKWYNWNEKRRVYEE

Signal sequence:

amino acids 1-34

Downloaded from www.elsevier.com

FIGURE 150

GCCCCAGGGACTGCTATGGCTTCCTTTGTTGTTACCCCCGGTCTGCGTCATGTTAAACTCCAATGTCTCTCTGTG
GTTAACTGCTCTTGCCATCAAGTTCACCCCTCATTGACAGCCAAGCACAGTATCCAGTTGTCAACACAAATTATGG
CAAAATCCGGGGCCTAAGAACACCGTTACCCAATGAGATCTTGGGTCCAGTGGAGCAGTACTTAGGGGTCCCCTA
TGCTCACCCCCCACTGGAGAGAGGCGGTTTTAGCCCCCAGAACCCCCGTCTCTCTGGACTGGCATCCGAAATAC
TACTCAGTTTGCTGCTGTGTGCCCCAGCACCTGGATGAGAGATCCTTACTGCATGACATGCTGCCCATCTGGTT
TACCGCCAATTTGGATACTTTGATGACCTATGTTCAAGATCAAAATGAAGACTGCCTTTACTTTAAACATCTACGT
GCCCCAGGAAGATGGAGCCAACACAAAGAAAAACGCAGATGATATAACGAGTAATGACCGTGGTGAAGACGAAGA
TATTCATGATCAGAACAGTAAGAAGCCCGTCATGGTCTATATCCATGGGGGATCTTACATGGAGGGCACCGGCAA
CATGATTGACGGCAGCATTTTTGGCAAGCTACGGAACGTCATCGTGATCACCATTAACTACCGTCTGGGAATACT
AGGGTTTTTAAGTACCGGTGACCAGGCAGCAAAAGGCAACTATGGGCTCCTGGATCAGATTCAAGCACTGCGGTG
GATTGAGGAGAATGTGGGAGCCTTTGGCGGGGACCCCAAGAGAGTGACCATCTTTGGCTCGGGGGCTGGGGCCTC
CTGTGTGACGCTGTTGACCCCTGTCCCCTACTCAGAAGGTCTCTTCCAGAAGGCCATCATTAGAGCGGCACCGC
CCTGTCCAGCTGGGCAGTGAATACCAGCCGGCCAAGTACACTCGGATATTGGCAGACAAGGTGCGGCTGCAACAT
GCTGGACACCACGGACATGGTAGAATGCCTGCGGAACAAGAACTACAAGGAGCTCATCCAGCAGACCATCACCCC
GGCCACCTACCACATAGCCTTCGGGCGGTGATCGACGGCGACGTCTCCAGACGACCCCCAGATCCTGATGGA
GCAAGGCGAGTTCTCTCAACTACGACATCATGCTGGGCGTCAACCAAGGGGAAGGCCCTGAAGTTCTGGGACGGCAT
CGTGGATAACGAGGACGGTGTGACGCCAACGACTTTGACTTCTCCGTGTCCAACCTTCGTGGACAACCTTTACGG
CTACCCTGAAGGGAAGACACTTTGCGGGAGACTATCAAGTTCTATGTACACAGACTGGGCCGATAAGGAAAACCC
GGAGACGCGCGGAAAACCTGCTGGTCTCTTTACTGACCACAGTGGGTGGCCCCCGCGGTGGCCGCGGACCT
GCACGCGCAGTACGGCTCCCCCACCCTACTTCTATGCCCTCTATCATCACTGCCAAAGCGAAATGAAGCCCAGCTG
GGCAGATTGGGCCCATGGTGTAGGTTCCCTATGTCTTCCGATCCCATGATCGGTCCCACCGAGCTCTTCAG
TTGTAACTTTTCCAAGAACGACGTCATGCTCAGCGCCGTGGTTCATGACCTACTGGACGAACTTCGCCAAAACCTG
TGATCCAAATCAACCAGTTCCTCAGGATACCAAGTTCATTACACAAAACCCACCGCTTTGAAGAAGTGGCCTG
GTCCAAGTATAATCCCAAAGACCAGCTCTATCTGCATATTGGCTTGAAACCCAGAGTGAGAGATCACTACCGGGC
AACGAAAGTGGCTTTCTGGTTGGAACTCGTTCCCTCATTTGCACAACCTTGAAACGAGATATTCCAGTATGTTTCAAC
AACCACAAAGGTTCTCCACCAGACATGACATCATTTCCCTATGGCACCCGCGATCTCCCGCCAAGATATGGCC
AACCACCAAACGCCCAGCAATCACTCTGCCAACAATCCCAAACACTCTAAGGACCCCTACAAAACAGGGCCTGA
GGACACAACCTGTCTCTATTGAAACCAAACGAGATTATTCCACCGAATTAAGTGTCAACATTGCCGTGGGGCGTC
GCTCCTCTTCTCAACATCTTAGCTTTTGGCGCGCTGTACTACAAAAGGACAAGAGGCGCCATGAGACTCACAG
GCGCCCCAGTCCCCAGAGAAACACCACAAATGATATCGCTCACATCCAGAACGAAGAGATCATGTCTCTGCAGAT
GAAGCAGCTGGAACACGATCACGAGTGTGAGTCGCTGCAGGCACACGACACACTGAGGCTCACCTGCCCGCCAGA
CTACACCCTCACGCTGCGCCGGTCGCCAGATGACATCCCACTTATGACGCCAAACACCATCACCATGATTCCAAA
CACACTGACGGGGATGCAGCCTTTGCACACTTTTAACACCTTCAGTGGAGGACAAAACAGTACAAATTTACCCCA
CGGACATTCCACCCTAGAGTATAGCTTTGCCCTATTTCCCTTCTATCCCTCTGCCCTACCCGCTCAGCAACAT
AGAAGAGGGAAGGAAAGAGAGAAGGAAAGAGAGAGAGAAAGAAAGTCTCCAGACCAGGAATGTTTTTGTCCCACT
GACTTAAGACAAAAATGCAAAAAGGCAGTCATCCCATCCCGGCAGACCCCTTATCGTTGGTGTTTTCCAGTATTAC
AAGATCAACTTCTGACCCTGTGAAATGTGAGAAGTACACATTTCTGTAAAATAACTGCTTTAAGATCTCTACCA
CTCCAATCAATGTTTAGTGTGATAGGACATCACCATTTCAAGGCCCGGGTGTTCACACGTATGGAAGCAGCT
GACACTTCTGAAACTCAGCCAAGGACACTTGATATTTTTTAATTACAATGGAAGTTTAAACATTTCTTTCTGTGC
CACACAATGGATGGCTCTCCTTAAGTGAAGAAAGAGTCAATGAGATTTTGCCAGCACATGGAGCTGTAATCCAG
AGAGAAGGAAACGTAGAAATTTATTATTAAGAAGTGGACTGTGCAGCGAAATCTGTACGGTTCTGTGCAAGAG
GTGTTTTGCCAGCCTGAACTATATTTAAGAGACTTTGT

FIGURE 151

MLNSNVLLWLTALAIKFTLIDSQAQYPVVNTNYGKIRGLRTPLPNEILGPVEQYLGVPYASP
PTGERRFQPPEPPSSWTGIRNTTQFAAVCPQHLDERSLLHDMLPIWFTANLDTLMTYVQDQN
EDCLYLNIYVPTEDGANTKKNADDITSNDRGEDEDIHDQNSKKPVMVYIHGGSYMEGTGNMI
DGSILASYGNVIVITINYRLGILGFLSTGDQAAKGNYGLLDQIQALRWIEENVGAFGGDPKR
VTIFGSGAGASCVSLLTLSHYSEGLFQKAI IQSGTALSSWAVNYQPAKYTRILADKVGCMNL
DTTDMVECLRNKNYKELIQQTITPATYHIAFGPVIDGDVIPDDPQILMEQGEFLNYDIMLG
VNOGEGLKFVDGIVDNEDGVTPNDFDFSVSNFVDNLYGYPEGKDTLRETIKFMYTDWADKENP
ETRRKTLVALFTDHQWVAPAVAADLHAQYGSPTYFYAFYHHCQSEMKPSWADSAHGDEV
PYVFGIPMIGPTELFSCNFSKNDVMLS AVVMTYWTNFAKTGDPNQVPQDTKFIHTKPNR
FEEVASKYNPKDQLYLHIGLKPRVRDHYRATKVAFWLELVPHLHNLNEIFQYVSTTTKV
PPDMSFPYGTRRSPAKIWPTTKRPAITPANNPKHSDPHKTGPEDTTVLIETKRDY
STELSVTIAVGASLLFLNILAFAALYYKKDKRRHETHRRPSPQRNTTNDIAHIQNEE
IMSLQMKQLEHDHECESLQAHDTLRLTCPPDYTLTLRRSPDDIPLMTPNTITMIPNT
LTGTMQPLHTFNTFSGGQNSTNLPHGHSTTRV

Signal sequence:

amino acids 1-24

Transmembrane domains:

amino acids 189-204, 675-692

Year	Age	Sex	Location	Species	Number	Percentage	Notes
1981	10-14	M	Urban	Urban	100	100%	
1982	10-14	M	Urban	Urban	100	100%	
1983	10-14	M	Urban	Urban	100	100%	
1984	10-14	M	Urban	Urban	100	100%	
1985	10-14	M	Urban	Urban	100	100%	
1986	10-14	M	Urban	Urban	100	100%	
1987	10-14	M	Urban	Urban	100	100%	
1988	10-14	M	Urban	Urban	100	100%	
1989	10-14	M	Urban	Urban	100	100%	
1990	10-14	M	Urban	Urban	100	100%	
1991	10-14	M	Urban	Urban	100	100%	
1992	10-14	M	Urban	Urban	100	100%	
1993	10-14	M	Urban	Urban	100	100%	
1994	10-14	M	Urban	Urban	100	100%	
1995	10-14	M	Urban	Urban	100	100%	
1996	10-14	M	Urban	Urban	100	100%	
1997	10-14	M	Urban	Urban	100	100%	
1998	10-14	M	Urban	Urban	100	100%	
1999	10-14	M	Urban	Urban	100	100%	
2000	10-14	M	Urban	Urban	100	100%	
2001	10-14	M	Urban	Urban	100	100%	
2002	10-14	M	Urban	Urban	100	100%	
2003	10-14	M	Urban	Urban	100	100%	
2004	10-14	M	Urban	Urban	100	100%	
2005	10-14	M	Urban	Urban	100	100%	
2006	10-14	M	Urban	Urban	100	100%	
2007	10-14	M	Urban	Urban	100	100%	
2008	10-14	M	Urban	Urban	100	100%	
2009	10-14	M	Urban	Urban	100	100%	
2010	10-14	M	Urban	Urban	100	100%	
2011	10-14	M	Urban	Urban	100	100%	
2012	10-14	M	Urban	Urban	100	100%	
2013	10-14	M	Urban	Urban	100	100%	
2014	10-14	M	Urban	Urban	100	100%	
2015	10-14	M	Urban	Urban	100	100%	
2016	10-14	M	Urban	Urban	100	100%	
2017	10-14	M	Urban	Urban	100	100%	
2018	10-14	M	Urban	Urban	100	100%	
2019	10-14	M	Urban	Urban	100	100%	
2020	10-14	M	Urban	Urban	100	100%	
2021	10-14	M	Urban	Urban	100	100%	
2022	10-14	M	Urban	Urban	100	100%	
2023	10-14	M	Urban	Urban	100	100%	
2024	10-14	M	Urban	Urban	100	100%	
2025	10-14	M	Urban	Urban	100	100%	
2026	10-14	M	Urban	Urban	100	100%	
2027	10-14	M	Urban	Urban	100	100%	
2028	10-14	M	Urban	Urban	100	100%	
2029	10-14	M	Urban	Urban	100	100%	
2030	10-14	M	Urban	Urban	100	100%	
2031	10-14	M	Urban	Urban	100	100%	
2032	10-14	M	Urban	Urban	100	100%	
2033	10-14	M	Urban	Urban	100	100%	
2034	10-14	M	Urban	Urban	100	100%	
2035	10-14	M	Urban	Urban	100	100%	
2036	10-14	M	Urban	Urban	100	100%	
2037	10-14	M	Urban	Urban</			

[illegible]

FIGURE 153

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA50911

><subunit 1 of 1, 348 aa, 1 stop

><MW: 39711, pI: 8.70, NX(S/T): 1

MAATLGPLGSWQQWRRCLSARDGSRMLLLLLLLGSGQGPQQVGAGQTFEYLKREHSLSKPYQ
GVGTGSSSLWNLMGNAMVMTQYIRLTPDMQSKQGALWNRVPCFLRDWELQVHFKEHGQKKN
LHGDGLAIWYTKDRMQPGPVFGNMDKFVGLGVFVDTPNEEKQQERVFPYISAMVNNGSLSY
DHERDGRPTELGGCTAIVRNHLYDTFLVIRYVKRHLTIMMDIDGKHEWRDCIEVPGVRLPRG
YYFGTSSITGDLSDNHDVISLKLFEFTVERTPEEEKLHRDVFLPSVDNMKLPMTAPLPPLS
GLALFLIVFFSLVFSVFAIVIGIILYNKWQEQSRKRFY

Signal sequence:

amino acids 1-38

Transmembrane domain:

amino acids 310-329

FIGURE 154

CCGAGCCGGGCGCGCAGCGACGGAGCTGGGGCCGGCCTGGGACCATGGGCGTGAGTGCAATCTACGGATCAGTCT
CTGATGGTGGGTGCTTAACCTCAGTGGGGACTCCAAGATTTCCATGAAGAAAATCAGTTGTCTTCATTCAAGAAT
TGGGGTCTGGCTCAGAATTCCTGCAGCTGGTGAAAATCTGTTTTCTAGAAGAGGTTAATTAATGCCTGCAGTCT
GACATGTTCCCGATTTGAGGTGAAACCATGAAGAGAAAATAGAATACTTAATAATGCTTTTCCGCAACCGCTTCT
TGCTGCTGCTGGCCCTGGCTGCGCTGCTGGCCTTTGTGAGCCTCAGCCTGCAGTTCTTCCACCTGATCCCGGTGT
CGACTCCTAAGAATGGAATGAGTAGCAAGAGTCGAAAGAGAATCATGCCCGACCTGTGACGGAGCCCCCTGTGA
CAGACCCCGTTTATGAAGCTCTTTTGTACTGCAACATCCCCAGTGTGGCCGAGCGCAGCATGGAAGGTGATGCCC
CGCATCATTTTAAAGCTGGTCTCAGTGCATGTGTTTCAATTCGCCACGGAGACAGGTACCCACTGTATGTCAATTCCTCA
AAACAAAGCGACCCAGAAATTGACTGCACCTCTGGTGGCTAACAGGAAACCGTATCACCCAAAACCTGGAAGCTTTCA
TTAGTCACATGTCAAAGGATCCGGAGCCTCTTTGAAAGCCCCCTGAACTCCTTGCCCTTTTACCCAAATCACC
CATTGTGTGAGATGGGAGAGCTCACACAGACAGGAGTTGTGCAGCATTTGCAGAACGGTCAGCTGCTGAGGGATA
TCTATCTAAAGAAACACAACTCCTGCCCAATGATTGGTCTGCAGACCAGCTCTATTTAGAGACCAGTGGGAAAA
GCCGGACCCCTACAAAGTGGGCTGGCCTTGCTTTATGGCTTTCTCCAGATTTTGACTGGAAGAAGATTTATTTCA
GGCACCAGCCAAGTGCAGTGTCTGCTCTGGAAGCTGCTATTGCCCGGTAAAGAAACAGTATCTGGAAGAGAGC
AGCGTCGTGAGTACCTCCTACGTTTGA AAAACAGCCAGCTGGAGAAGACCTACGGGAGATGGCCAAGATCGTGG
ATGTCCCCACCAAGCAGCTTAGAGCTGCCAACCCCATAGACTCCATGCTCTGCCACTTCTGCCACAATGTGAGCT
TTCCCTGTACCAGAAATGGCTGTGTTGACATGGAGCACTTCAAGGTAATTAAGACCCATCAGATCGAGGATGAAA
GGGAAAGACGGGAGAAGAAATTGTACTTCGGGTATTTCTCTCTGGGTGCCACCCCATCTGAACCAAACCATCG
GCCGGATGCAGCGTGCCACCGAGGGCAGGAAAGAAGAGCTCTTGCCCTCTACTCTGCTCATGATGTCACTCTGT
CACCAGTTCTCAGTGCCTTTGGGCCCTTTCAGAAAGCCAGGTTCCCAAGGTTTGCAGCCAGGTTGATCTTTGAGCTTT
GGCAAGACAGAGAAAAGCCCAAGTGAACATTCGGTCCGGATTCTTTACAATGGCGTCGATGTACATTCACACCT
CTTTCTGCCAAGACCACCACAAGCGTTCTCCCAAGCCCATGTGCCCGCTTGAAAACCTTGGTCCGCTTTGTGAAAA
GGGACATGTTTGTAGCCCTGGGTGGCAGTGGTACAAATTATTATGATGCATGTACAGGGAAGGATTCTAAAAGG
TATGCAGTACAGCAGTATAGAATCCATGCCAATACAGAGCATAGGGAAAGGTCCACTTCTAGTTTGTCTGTTAC
TAAGGGTAGAAGATTATTGCTTTTAAAGGCTAAATATTGTTTGTGGGAACACAGATGGTTGGGGTTGAACAGT
AAGCACATTGCTGCAATGTGGTACGTGAATTGCTTGGTACAAAATGGCCAGTTTACAGAGGAATAGAAGGTACTT
TATCATAGCCAGACTTCGCTTAGAATGCCAGAATAATATAGTTCAAGACCTGAAGTTGCCAATCCAAGTTGCAC
TCTTCTGGCCTGCCCATGTTACTATGTGATGGAACCAGCACACCTCAACCAAAATTTTTTAACTTTAGACATT
TTTACCTTGTCTTGTTAAGAATTTCTTGAAGTGATTATCTAAAATAAAGGTTGGCAAACCTTTTTCTGTAAAGG
GCCAGATTGTAAATATTTTCAGACTGTGTGGACCAAAAGGCCACATACAGTCTCTGTGCATACTACTCAACTCTGT
TTCTGAAGCAGGAAAGCCACCACAGACAGTACATAAAGGAATATGTGTAGCTGGGTTCCAGGCCAGACAAAACA
GATGGTGACCAGACTTGGCCCTGGGCTGTAGTTTGTGACCCCTCATCTAAAAATAGGCTATACTACAATTGC
ACTTCCAGCACTTTGAGAACGAGTTGAATACCAAGAATTATTCAATGGTTTCTCCAGTAACCTCTGCTAGAAACA
CAGAATTTGGTCTGTATCTGACACTAGAACAAAACCTTGAGGGTAAATAAATGAATTAGAATGAATCATAGAA
AACTGATTAGAAGAATACCTGATGTTTATGATGATTGTGGTACAAGATAGTTTTAAGTATGTTCTAAATATTTGT
CTGCTGTAGTCTATTTGCTGTATATGCTGAAATTTTTGTATGCCATTTAGTATTTTTATAGTTTAGGAAAATATT
TTCTAAGACCAGTTTATAGTACTCTTATTCTGTAGTAATATTCAATTTGCTGTACCTGCTTGGTGGTTAGAAG
GAGGCTAGAAGATGAATTCAGGCACTTTCTTCCAATAAACTAATTATGGCTCATTCCCTTTGACAAGCTGTAGA
ACTGGATTCATTTTTAAACCATTTTCATCAGTTTCAAATGGTAAATTCTGATTGATTTTTTAAATGCGTTTTTGGGA
AGAACTTTGCTATTAGGTAGTTTACAGATCTTTATAAGGTGTTTTATATATTAGAAGCAATTATAATTACATCTG
TGATTTCTGAACTAATGGTGCTAATTACAGAGAAATGGAAAGTGAAAGTGAGATTCTCTGTTGTGATCGGCATTCC
AACTTTTTCTCTTTGTTTTTGTCCAGTGTGCAATTTGAATATGTCTGTTTCTATAAATAAATTTTTTAAAGAATAA

FIGURE 155

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48329

><subunit 1 of 1, 480 aa, 1 stop

><MW: 55240, pI: 9.30, NX(S/T): 2

MLFRNRFLLLLALAALLAFVSLSLQFFHLIPVSTPKNGMSSKSRKRIMPDVTEPPVTDVY
EALLYCNIPPSVAERSMEGHAPHHFKLVSVHVFIRHGDRYPLYVIPKTKRPEIDCTLVANRKP
YHPKLEAFISHMSKSGSGASFESPLNSLPLYPNHPLCEMGELTQTGVVQHLQNGQLLRDIYLK
KHKLLPNDWSADQLYLETTGKSRTLQSGLALLYGFLPDFDWKKIYFRHQPSALFCSGSCYCP
VRNQYLEKEQRRQYLLRLKNSQLEKTYGEMAKIVDVPTKQLRAANPIDSMMLCHFCHNVSFPC
TRNGCVDMEHFKVIKTHQIEDERERREKKLYFGYSLLGAHPILNQTIGRMQRATEGRKEELF
ALYSAHDVTLSPVLSALGLSEARFPRFAARLIFELWQDREKPSSEHSVRILYNGVDVTFHTSF
CQDHHKRSPKPMCPLNLVRFVKRDMFVALGGSGTNYDACHREGF

Signal sequence:

amino acids 1-18

FIGURE 155

FIGURE 157

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48306

><subunit 1 of 1, 916 aa, 1 stop

><MW: 100204, pI: 4.92, NX(S/T): 4

MIPARLHRDYKGLVLLGILLGTLWETGCTQIRYSVPPEELEKGSRVGDISRDLGLEPRELAER
GVRIIPRGRTQLFALNPRSGSLVTAGRIDREELCMGAIKCQLNLDILMEDKVKIYGVEVEVR
DINDNAPYFRESELEIKISENAATEMRFPPLHAWDPDIGKNSLQSYELSPNTHFSLIVQNGA
DGSKYPELVVKRALDREEKAAHHLVLTASDGGDPVRTGTARIRVMVLDANDNAPAFAPQPEYR
ASVPENLALGTQLLVVNATDPDEGVNAEVRYSFYVDDKAAQVFKLDCNSGTISTIGELDHE
ESGFYQMEVQAMDNAGYSARAKVLITVLDVNDNAPEVVLTSASSVPENSPRGTLIALLVN
DQDSEENGQVICFIQGNLPFKLEKSYGNYYSLVTDIVLDREQVPSYNITVTATDRGTPPLST
ETHISLVNADTNDNPPVFPQASYSAYIPENNPRGVSLVSVTAHDPDCEENAQITYSLAENTI
QGASLSSYVSINS DTGVLIALSSFDYEQFRDLQVKVMARDNGHPPLSSNVLSLFLVLDQNDN
APEILYPALPTDGSTGVELAPRSAEPGYLVTKVAVDRDSGQNAWLSYRLLKASEPGLFSVG
LHTGEVRTARALLDRDALKQSLVVAVQDHGQPPLSATVTTLTVAVADSIPQVLADLGSLES
NSETSDLTLYLVVAVAAVSCVFLAFVILLALLRLRRWHKSRLQASGGGLTGAPASHFVGVD
GVQAFLLQYTSHEVSLTTDSRKSHLIFPQPNYADMLVSQESFEKSEPLLLSGDSVFSKDSHGL
IEVSLYQIFFLFFFNCSVSQAGVQRYDHSSLRPQTPRLKQLSHLCLRCNRDYZRCKPPTVCLS
IYLSIYLSIYLSIYLLLSCTDGS LTPVIPVLWEAEAGGSPEVGSLRPA

Signal sequence:

amino acids 1-30

Transmembrane domains:

amino acids 693-711, 809-823, 869-888

FIGURE 158

CCCAGGCTCTAGTGCAGGAGGAGAAGGAGGAGGAGCAGGAGGTGGAGATTCCCAGTTAAAAG
GCTCCAGAATCGTGTACCAGGCAGAGAACTGAAGTACTGGGGCCTCCTCCACTGGGTCCGAA
TCAGTAGGTGACCCCGCCCCTGGATTCTGGAAGACCTCACCATGGGACGCCCCCGACCTCGT
GCGGCCAAGACGTGGATGTTCTTGCTCTTGCTGGGGGAGCCTGGGCAGGACACTCCAGGGC
ACAGGAGGACAAGGTGCTGGGGGGTTCATGAGTGCCAACCCCATTCGCAGCCTTGGCAGGCGG
CCTTGTTCCAGGGCCAGCAACTACTCTGTGGCGGTGTCCTTGTAGGTGGCAACTGGGTCCTT
ACAGCTGCCCCTGTAAAAAACCGAAATACACAGTACGCCTGGGAGACCACAGCCTACAGAA
TAAAGATGGCCCAGAGCAAGAAATACCTGTGGTTCAGTCCATCCCACACCCCTGCTACAACA
GCAGCGATGTGGAGGACCACAACCATGATCTGATGCTTCTTCAACTGCGTGACCAGGCATCC
CTGGGGTCCAAAGTGAAGCCCATCAGCCTGGCAGATCATTGCACCCAGCCTGGCCAGAAGTG
CACCGTCTCAGGCTGGGGCACTGTCACCAGTCCCCGAGAGAATTTTCCTGACACTCTCAACT
GTGCAGAAGTAAAAATCTTTCCCCAGAAGAAGTGTGAGGATGCTTACCCGGGGCAGATCACA
GATGGCATGGTCTGTGCAGGCAGCAGCAAAGGGGCTGACACGTGCCAGGGCGATTCTGGAGG
CCCCCTGGTGTGTGATGGTGCACCTCCAGGGCATCACATCCTGGGGCTCAGACCCCTGTGGGA
GGTCCGACAAACCTGGCGTCTATACCAACATCTGCCGCTACCTGGACTGGATCAAGAAGATC
ATAGGCAGCAAGGGCTGATTCTAGGATAAGCACTAGATCTCCCTTAATAAACTCACAACTCT
CTGGTTC

FIGURE 159

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48336

<subunit 1 of 1, 260 aa, 1 stop

<MW: 28048, pI: 7.87, NX(S/T): 1

MGRPRPRAAKTWMFLLLLGGAWAGHSRAQEDKVLGGHECQPHSQPWQAALFQGQQLLCGGVL
VGGNWVLTAAHCKKPKYTVRLGDHSLQNKDGPEQEIPVVQSIPHPCYNSSDVEDHNHDLMLL
QLRDQASLGSKVKPISLADHCTQPGQKCTVSGWGTVTSPRENFPDTLNCAEVKIFPQKKCED
AYPGQITDGMVCAGSSKGADTCQGDSGGPLVCDGALQGITSWGS DPCGRSDKPGVYTNICRY
LDWIKKIIGSKG

Important Features:

Signal peptide:

amino acids 1-23

Transmembrane domain:

amino acids 51-71

N-glycosylation site.

amino acids 110-113

Serine proteases, trypsin family, histidine active site.

amino acids 69-74 and 207-217

Tyrosine kinase phosphorylation site.

amino acids 182-188

Kringle domain proteins motif

amino acids 205-217

FIGURE 160

GGCGCCGGTGCACCGGGCGGGCTGAGCGCCTCCTGCGGCCCGGCCTGCGCGCCCCGGCCCCG
CGCGCCGCCACGCCCCAACCCCGGCCCGCGCCCCCTAGCCCCCGCCCGGGCCCGCGCCCCG
GCCCGCGCCAGGTGAGCGCTCCGCCCCGCGGAGGCCCGCCCGGCCCGCCCCCGCCCCG
CCCCGGCCGGCGGGGAACCGGGCGGATTCTCGCGCGTCAAACCACCTGATCCCATAAAAC
ATTCATCCTCCCGGCGGCCCGCGCTGCGAGCGCCCCGCCAGTCCGCGCCGCGCCGCCCCCTCG
CCCTGTGCGCCCTGCGCGCCCTGCGCACCCGCGGCCCGAGCCAGCCAGAGCCGGGCGGAGC
GGAGCGCGCCGAGCCTCGTCCCGCGGCCCGGGCCGGGGCCGGGCCGTAGCGGCGGCGCCTGGA
TGCGGACCCGGCCGCGGGGAGACGGGCGCCCGCCCCGAAACGACTTTCAGTCCCCGACGCGC
CCCGCCCAACCCCTACGATGAAGAGGGCGTCCGCTGGAGGGAGCCGGCTGCTGGCATGGGTG
CTGTGGCTGCAGGCCTGGCAGGTGGCAGCCCCATGCCAGGTGCCTGCGTATGCTACAATGA
GCCCAAGGTGACGACAAGCTGCCCCCAGCAGGGCCTGCAGGCTGTGCCGTGGGCATCCCTG
CTGCCAGCCAGCGCATCTTCTGACAGGCAACCGCATCTCGCATGTGCCAGCTGCCAGCTTC
CGTGCCTGCCGCAACCTCACCATCCTGTGGCTGCACTCGAATGTGCTGGCCCGAATTGATGC
GGCTGCCTTCACTGGCCTGGCCCTCCTGGAGCAGCTGGACCTCAGCGATAATGCACAGCTCC
GGTCTGTGGACCTGCCACATTCCACGGCCTGGGCCGCCTACACACGCTGCACCTGGACCGC
TGCGGCCTGCAGGAGCTGGGCCCGGGGCTGTTCCGCGGCCTGGCTGCCCTGCAGTACCTCTA
CCTGCAGGACAACGCGCTGCAGGCACTGCCTGATGACACCTTCCGCGACCTGGGCAACCTCA
CACACCTCTTCTGACGGCAACCGCATCTCCAGCGTGCCCGAGCGCGCCTTCCGTGGGCTG
CACAGCCTCGACCGTCTCCTACTGCACCAGAACCGCGTGGCCCATGTGCACCCGCATGCCTT
CCGTGACCTTGGCCGCCTCATGACACTCTATCTGTTTGCCAACAATCTATCAGCGCTGCCCA
CTGAGGCCCTGGCCCCCTGCGTGCCCTGCAGTACCTGAGGCTCAACGACAACCCCTGGGTG
TGTGACTGCCGGGCACGCCCCTCTGGGCCTGGCTGCAGAAGTTCCGCGGCTCCTCCTCCGA
GGTGCCCTGCAGCCTCCCGCAACGCCTGGCTGGCCGTGACCTCAAACGCCTAGCTGCCAATG
ACCTGCAGGGCTGCGCTGTGGCCACCGGCCCTTACCATCCCATCTGGACCGGCAGGGCCACC
GATGAGGAGCCGCTGGGGCTTCCCAAGTGCTGCCAGCCAGATGCCGCTGACAAGGCCTCAGT
ACTGGAGCCTGGAAGACCAGCTTCGGCAGGCAATGCGCTGAAGGGACGCGTGCCGCCCGGTG
ACAGCCCGCCGGGCAACGGCTCTGGCCACGGCACATCAATGACTCACCCTTTGGGACTCTG
CCTGGCTCTGCTGAGCCCCCGCTCACTGCAGTGCGGCCCGAGGGCTCCGAGCCACCAGGGTT
CCCCACCTCGGGCCCTCGCCGGAGGCCAGGCTGTTACGCAAGAACCGCACCCCGCAGCCACT
GCCGTCTGGGCCAGGCAGGCAGCGGGGTGGCGGGACTGGTGACTCAGAAGGCTCAGGTGCC
CTACCCAGCCTCACCTGCAGCCTCACCCCCCTGGGCCTGGCGCTGGTGCTGTGGACAGTGCT
TGGGCCCTGCTGAACCCCCAGCGGACACAAGAGCGTGCTCAGCAGCCAGGTGTGTGTACATAC
GGGGTCTCTCTCCACGCCGCCAAGCCAGCCGGGCGGCCGACCCGTGGGGCAGGCCAGGCCAG
GTCTCTCCCTGATGGACGCCTGCCGCCCGCCACCCCCATCTCCACCCCATCATGTTTACAGGG
TTCGGCGGCAGCGTTTGTTCCAGAACGCCGCCTCCACCCAGATCGCGGTATATAGAGATAT
GCATTTTATTTTACTTGTGTAAAAATATCGGACGACGTGGAATAAAGAGCTCTTTTCTTAA
AAAA

FIGURE 161

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA44184

><subunit 1 of 1, 473 aa, 1 stop

><MW: 50708, pI: 9.28, NX(S/T): 6

MKRASAGGSRL LAWVLWLQAWQVAAPCPGACVCYNEPKVTTSCPQQGLQAVPVGIPAASQRI
FLHGNRISHVPAASFRACRNLTILWLHSNVLARIDAAAF TGLALLEQLDLSDNAQLRSVDPA
TFHGLGRLHTLHLDRCGLQELGPG LFRGLAALQYLYLQDNALQALPDDTFRDLGNLTHLFLH
GNRISSVPERAFRGLHSLDRLLLLHQNRVAHVHPHAFRDLGRLMTLYLFANNLSALPTEALAP
LRALQYLRRLNDNPWVCDRCRARPLWAWLQKFRGSSSEVPCSLPQRLAGRDLKRLAANDLQGCA
VATGPYHPIWTGRATDEEPLGLPKCCQPDAA DKASVLEPGRPASAGNALKGRVPPGDSPPGN
GSGPRHINDSPFGTLPGSAEPPLTAVRPEGSEPPGFPTSGPRRRPGCSRKNRTRSHCRLGQA
GSGGGGTGDSEGGALPSLTCSLTPLGLALVLWTVLGPC

Important features:

Signal peptide:

amino acids 1-26

Leucine zipper pattern.

amino acids 135-156

Glycosaminoglycan attachment site.

amino acids 436-439

N-glycosylation site.

amino acids 82-85, 179-183, 237-240, 372-375 and 423-426

VWFC domain

amino acids 411-425

[illegible]

GGAAGTCCACGGGGAGCTTGGATGCAAGAGGGAGGACGGCTCCTCTGGAGAGGACTAC
 TCACTGGCATATTTCTGAGGTATCTGTAGAATAACACAGCCTCAGATACTGGGGACTTTAC
 AGTCCCACAGAACCGTCTCTCCAGGAAGCTGAATCCAGCAAGAACAATGGAGGCCAGCGGA
 AGCTCATTTGCAGACAAAGGCAAGTCTCTTTTTTCTCTTTCTCTTTTGGGCTTATCTCTGGCG
 GGCGCGGCGGAACCTAGAAGCTATTCTGTGGTGGAGGAACTGAGGGCAGCTCCTTTGTAC
 CAATTTAGCAAAGGACCTGGGTCTGGAGCAGAGGGAATTCTCCAGGCGGGGGGTTAGGGTTG
 TTTCCAGAGGGAACAAACTACATTTGCAGCTCAATCAGGAGACCGCGGATTGTGTGCTAAAT
 GAGAAATTGGACCGTGAGGATCTGTGCGGTACACAGAGCCCTGTGTGCTACGTTTCCAAG
 GTTGCTAGAGAGTCCCTTCGAGTTTTTTCAAGCTGAGCTGCAAGTAATAGACATAAACGACC
 ACTCTCCAGTATTTCTGGACAAACAAATGTTGGTGAAAGTATCAGAGAGCAGTCTCTCTGGG
 ACTACGTTTCTCTGAAGAATGCCGAAGACTTAGATGTAGGCCAAAACAATATTGAGAACTA
 TATAATCAGCCCCAACTCCTATTTTCGGGTCTCACCCGCAAACGCAGTGATGGCAGGAAAT
 ACCCAGAGCTGGTGCTGGACAAAGCGCTGGACCGAGAGGAAGAAGCTGAGCTCAGGTTAACA
 CTCACAGCACTGGATGGTGGCTCTCCGCCAGATCTGGCACTGCTCAGGTCTACATCGAAGT
 CCTGGATGTCAACGATAATGCCCTGAATTTGAGCAGCCTTTCTATAGAGTGCAGATCTCTG
 AGGACAGTCCGGTAGGCTTCTGTTGTGAAGGTCTCTGCCACGGATGTAGACACAGGAGTC
 AACGGAGAGATTTCTTATTCATTTTTCCAAGCTTCAGAAGAGATTGGCAAACCTTTAAGAT
 CAATCCCTTGACAGGAGAAATTGAACTAAAAAACAACCTCGATTTGAAAAACTTCAGTCTT
 ATGAAGTCAATATTGAGGCAAGAGATGCTGGAACCTTTCTGGAAAATGCACCGTTCTGATT
 CAAGTGATAGATGTGAACGACCATGCCCCAGAAGTTACCATGTCTGCATTTACCAGCCCAAT
 ACCTGAGAACGCGCCTGAAACTGTGGTTGCACTTTTCAGTGTTTCAGATCTTGATTGAGGAG
 AAAATGGGAAAATTAGTTGCTCCATTGAGGAGGATCTACCCTTCTCTGAAATCCGCGGAA
 AACTTTTACACCCTACTAACGGAGAGACCACTAGACAGAGAAAGCAGAGCGGAATACAACAT
 CACTATCACTGTCACTGACTTGGGGACCCCTATGCTGATAACACAGCTCAATATGACCGTGC
 TGATCGCCGATGTCAATGACAACGCTCCCGCCTTCACCCAAACCTCCTACACCCTGTTTCGT
 CGCGAGAAACACAGCCCCGCGCTGCACATCCGCAGCGCTCAGCGCTACAGACAGAGACTCAGG
 CACCAACGCCCAGGTACCTACTCGCTGCTGCCGCCCCAGGACCCGCACCTGCCCTCACAT
 CCTGGTCTCCATCAACGCGGACAACGGCCACCTGTTTCGCCCTCAGGTCTCTGGACTACGAG
 GCCCTGCAGGGGTTCCAGTTCCGCGTGCGGCTTCAGACCAAGGCTCCCCGGCGCTGAGCAG
 CGAGGCGCTGGTGCGCGTGGTGGTGCTGGACGCCAACGACAACCTCGCCCTTCGTGCTGTACC
 CGCTGCAGAACGGCTCCGCGCCCTGCACCGAGCTGGTGCCCCGGGCGGCGAGCCGGGCTAC
 CTGGTGACCAAGGTGGTGGCGGTGGACGGCGACTCGGGCCAGAACGCCTGGCTGTCTACCA
 GCTGCTCAAGGCCACGGAGCTCGGTCTGTTTCGGCGTGTGGGCGCACAAATGGCGAGGTGCGCA
 CCGCCAGGCTGCTGAGCGAGCGCGACGCGGCCAAGCACAGGCTGGTGGTGCTGGTCAAGGAC
 AATGGCGAGCCTCCGCGCTCGGCCACCGCCACGCTGCACGTGCTCCTGGTGGACGGCTTCTC
 CCAGCCCTACCTGCCTCTCCCGGAGGCGGCCCGACCCAGGCCCAGGCCGACTTGCTCACCG
 TCTACCTGGTGGTGGCGTTGGCCTCGGTGTCTTCGCTCTTCTCTTTTCGGTGCTCCTGTTT
 GTGGCGGTGCGGCTGTGTAGGAGGAGCAGGGCGGCTCGGTGGGTGCTGCTTGGTGCCCGA
 GGGCCCCCTTCCAGGGCATCTTGTGGACATGAGCGGCACCAGGACCTTATCCAGAGCTACC
 AGTATGAGGTGTGTCTGGCAGGAGGCTCAGGGACCAATGAGTTCAAGTTCCTGAAGCCGATT
 ATCCCCAACTTCCCTCCCCAGTGCCCTGGGAAAGAAATACAAGGAAATCTACCTTCCCCAA
 TAACTTTGGGTTCAATATTCAGTGAACCATAGTTGACTTTTACATTCCATAGGTATTTTATTT
 TGTGGCATTTCCATGCCAATGTTTATTTCCCCCAATTTGTGTGTATGTAATATTGTACGGAT
 TTAATCTTGATTTTTCTCATGTTCTTTCTCCCTTTGTTTTAAAGTGAACATTTACCTTTATTT
 CCTGGTTCTT

FIGURE 163

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48314

<subunit 1 of 1, 798 aa, 1 stop

<MW: 87552, pI: 4.84, NX(S/T): 5

MEASGKLICRQRQVLFSFLLLGLSLAGAAEPRSYVVEETEGSSFVTNLAKDLGLEQREFSR
RGVRVVSARGNKLHLQLNQETADLLLNEKLDREDLCGHTEPCVLRQVLLESPFEFFQAELQV
IDINDHSPVFLDKQMLVKVSESSPPGTTFPLKNAEDLDVGQNNIENYIISPNSYFRVLTRKR
SDGRKYPELVLDKALDREEEAELRLTLTALDGGSPPRSGTAQVYIEVLDVNDNAPEFEQPFY
RVQISEDSPVGFLVVKVSATDVDTGNGEISYSLFQASEEIGKTFKINPLTGEIELKKQLDF
EKLQSYEVNIEARDAGTFSGKCTVLIQVIDVNDHAPEVTMSAFTSPIPENAPETVVALFSVS
DLDSGENGKISCSIQEDLPFLLKSAENFYTLTTERPLDRESRAEYNITITVTDLGTPLMLITQ
LNMTVLIADVNDNAPAFQTQSYTLFVRENNSPALHIRSVSATDRDSGTNAQVTYSLLPPQDP
HLPLTSLVSINADNGHLFALRSLDYEALQGFQFRVGASDHGSPALSSEALVRVVVLDANDNS
PFVLYPLQNGSAPCTELVPRAAEPGYLVTKVVAVDGDSGQNAWLSYQLLKATELGFLFGVWAH
NGEVRTARLLSERDAAKHRLVVLVKDNGEPPRSATATLHVLLVDGFSQPYLPLPEAAPTQAA
ADLLTVYLVVALASVSSLFLFSVLLFVAVRLCRRSRAASVGRCLVPEGPLPGHLVDMSGTRT
LSQSYQYEVCLAGGSGTNEFKFLKPIIPNFPPQCPGKEIQGNSTFPNNFGFNIQ

Important features:

Signal peptide:

amino acids 1-26

Transmembrane domain:

amino acids 685-712

Cadherins extracellular repeated domain signature.

amino acids 122-132, 231-241, 336-346, 439-449 and 549-559

ATP/GTP-binding site motif A (P-loop).

amino acids 285-292

N-glycosylation site.

amino acids 418-421, 436-439, 567-570 and 786-789

Table 1. Demographic characteristics of the study population	
Characteristic	Frequency (%)
Age (years)	
< 18	10 (10.0)
18-24	20 (20.0)
25-34	30 (30.0)
35-44	25 (25.0)
45-54	15 (15.0)
≥ 55	10 (10.0)
Sex	
Male	40 (40.0)
Female	60 (60.0)
Ethnicity	
White	30 (30.0)
Black	20 (20.0)
Hispanic	10 (10.0)
Other	40 (40.0)
Education level	
High school or less	20 (20.0)
Some college	30 (30.0)
Bachelor's degree	20 (20.0)
Master's degree	10 (10.0)
PhD	10 (10.0)
Income (USD/year)	
< 10,000	10 (10.0)
10,000-20,000	20 (20.0)
20,000-30,000	30 (30.0)
30,000-40,000	20 (20.0)
≥ 40,000	10 (10.0)
Marital status	
Married	30 (30.0)
Single	20 (20.0)
Divorced	10 (10.0)
Widowed	10 (10.0)
Never married	10 (10.0)
Health insurance	
Medicaid	10 (10.0)
Medicare	20 (20.0)
Private	30 (30.0)
Uninsured	10 (10.0)
Other	10 (10.0)

Table 1. Demographic characteristics of the study population	
Characteristic	Frequency (%)
Age (years)	
< 18	10 (10.0)
18-24	20 (20.0)
25-34	30 (30.0)
35-44	25 (25.0)
45-54	15 (15.0)
≥ 55	10 (10.0)
Sex	
Male	40 (40.0)
Female	60 (60.0)
Ethnicity	
White	30 (30.0)
Black	20 (20.0)
Hispanic	10 (10.0)
Other	40 (40.0)
Education level	
High school or less	20 (20.0)
Some college	30 (30.0)
Bachelor's degree	20 (20.0)
Master's degree	10 (10.0)
PhD	10 (10.0)
Marital status	
Married	30 (30.0)
Single	20 (20.0)
Divorced	10 (10.0)
Widowed	10 (10.0)
Never married	10 (10.0)
Employment status	
Employed	30 (30.0)
Unemployed	20 (20.0)
Retired	10 (10.0)
On disability	10 (10.0)
Homemaker	10 (10.0)
Student	10 (10.0)
Health insurance	
Medicare	10 (10.0)
Medicaid	20 (20.0)
Private	30 (30.0)
Uninsured	10 (10.0)
Other	10 (10.0)
Annual income	
< \$10,000	10 (10.0)
\$10,000-\$19,999	20 (20.0)
\$20,000-\$29,999	10 (10.0)
\$30,000-\$39,999	10 (10.0)
\$40,000-\$49,999	10 (10.0)
\$50,000-\$59,999	10 (10.0)
\$60,000-\$69,999	10 (10.0)
\$70,000-\$79,999	10 (10.0)
\$80,000-\$89,999	10 (10.0)
\$90,000-\$99,999	10 (10.0)
\$100,000+	10 (10.0)

FIGURE 166

CCCGGCTCCGCTCCCTCTGCCCCCTCGGGGTGCGCGCCCCACGATGCTGCAGGGCCCTGGCT
CGCTGCTGCTGCTCTTCCTCGCCTCGCACTGCTGCCTGGGCTCGGCGCGGGGCTCTTCCTC
TTTGGCCAGCCCGACTTCTCCTACAAGCGCAGCAATTGCAAGCCCATCCCGGTCAACCTGCA
GCTGTGCCACGGCATCGAATACCAGAACATGCGGCTGCCCAACCTGCTGGGCCACGAGACCA
TGAAGGAGGTGCTGGAGCAGGCCGGCGCTTGGATCCCGCTGGTCATGAAGCAGTGCCACCCG
GACACCAAGAAGTTCCTGTGCTCGCTCTTCGCCCCCGTCTGCCTCGATGACCTAGACGAGAC
CATCCAGCCATGCCACTCGCTCTGCGTGCAGGTGAAGGACCGCTGCGCCCCGGTCATGTCCG
CCTTCGGCTTCCCCTGGCCCCGACATGCTTGAGTGCGACCGTTTCCCCCAGGACAACGACCTT
TGCATCCCCCTCGCTAGCAGCGACCACCTCCTGCCAGCCACCGAGGAAGCTCCAAAGGTATG
TGAAGCCTGCAAAAATAAAAATGATGATGACAACGACATAATGGAAACGCTTTGTAAAAATG
ATTTTGCAGTGAAAATAAAAGTGAAGGAGATAACCTACATCAACCGAGATACCAAAATCATC
CTGGAGACCAAGAGCAAGACCATTTACAAGCTGAACGGTGTGTCCGAAAGGGACCTGAAGAA
ATCGGTGCTGTGGCTCAAAGACAGCTTGCAGTGCACCTGTGAGGAGATGAACGACATCAACG
CGCCCTATCTGGTCATGGGACAGAAACAGGGTGGGGAGCTGGTGATCACCTCGGTGAAGCGG
TGGCAGAAGGGGCAGAGAGAGTTCAAGCGCATCTCCCGCAGCATCCGCAAGCTGCAGTGCTTA
GTCCCGGCATCCTGATGGCTCCGACAGGCCTGCTCCAGAGCACGGCTGACCATTCTGCTCC
GGGATCTCAGCTCCCGTTCCCCAAGCACACTCCTAGCTGCTCCAGTCTCAGCCTGGGCAGCT
TCCCCCTGCCTTTTGCACGTTTGCATCCCCAGCATTTCTTGAGTTATAAGGCCACAGGAGTG
GATAGCTGTTTTACCTAAAGGAAAAGCCCACCCGAATCTTGTAGAAATATTCAAACCTAATA
AAATCATGAATATTTTAA

[illegible]

><MW: 33518, pI: 7.74, NX(S/T): 0

Important features:

amino acids 1-20

amino acids 6-153

FIGURE 168

GTGGAGGCCGCCGACGATGGCGGGGCCGACGGAGGCCGAGACGGGGTTGGCCGAGCCCCGGG
CCCTGTGCGCGCAGCGGGGCCACCGCACCTACGCGCGCCGCTGGGTGTTCCCTGCTCGCGATC
AGCCTGCTCAACTGCTCCAACGCCACGCTGTGGCTCAGCTTTGCACCTGTGGCTGACGTCAT
TGCTGAGGACTTGGTCCTGTCCATGGAGCAGATCAACTGGCTGTCACTGGTCTACCTCGTGG
TATCCACCCCATTTTGGCGTGGCGGCCATCTGGATCCTGGACTCCGTGCGGGCTCCGTGCGGCG
ACCATCCTGGGTGCGTGGCTGAACTTTGCCGGGAGTGTGCTACGCATGGTGCCCTGCATGGT
TGTTGGGACCCAAAACCCATTTGCCTTCCTCATGGGTGGCCAGAGCCTCTGTGCCCTTGCCC
AGAGCCTGGTCATCTTCTCTCCAGCCAAGCTGGCTGCCTTGTGGTTCCAGAGCACCAGCGA
GCCACGGCCAACATGCTCGCCACCATGTGCAACCCTCTGGGCGTCCTTGTGGCCAATGTGCT
GTCCCCCTGTGCTGGTCAAGAAGGGTGAGGACATTCCGTTAATGCTCGGTGTCTATACCATCC
CTGCTGGCGTTCGTCTGCCTGCTGTCCACCATCTGCCTGTGGGAGAGTGTGCCCCCACC
CCCTCTGCCGGGGCTGCCAGCTCCACCTCAGAGAAGTTCCCTGGATGGGCTCAAGCTGCAGCT
CATGTGGAACAAGGCCTATGTCTATCCTGGCTGTGTGCTTGGGGGAATGATCGGGATCTCTG
CCAGCTTCTCAGCCCTCCTGGAGCAGATCCTCTGTGCAAGCGGCCACTCCAGTGGGTTTTCC
GGCCTCTGTGGCGCTCTCTTCATCACGTTTGGGATCCTGGGGGCACTGGCTCTCGGCCCTTA
TGTGGACCGGACCAAGCACTTCACTGAGGCCACCAAGATTGGCCTGTGCCTGTTCTCTCTGG
CCTGCGTGCCCTTTGCCCTGGTGTCCAGCTGCAGGGACAGACCCTTGCCCTGGCTGCCACC
TGCTCGCTGCTCGGGCTGTTTGGCTTCTCGGTGGGCCCCGTGGCCATGGAGTTGGCGGTCTGA
GTGTTCCCTTCCCCGTGGGGGAGGGGGCTGCCACAGGCATGATCTTTGTGCTGGGGCAGGCCG
AGGGAATACTCATCATGCTGGCAATGACGGCACTGACTGTGCGACGCTCGGAGCCGTCCTTG
TCCACCTGCCAGCAGGGGGAGGATCCACTTGA CTGGACAGTGTCTCTGCTGCTGATGGCCGG
CCTGTGCACCTTCTTCAGCTGCATCCTGGCGGTCTTCTTCCACACCCCATACCGGCGCCTGC
AGGCCGAGTCTGGGGAGCCCCCTCCACCCGTAACGCCGTGGGCGGCGCAGACTCAGGGCCG
GGTGTGGACCGAGGGGGAGCAGGAAGGGCTGGGGTCTTGGGGCCAGCACGGCGACTCCGGA
GTGCACGGCGAGGGGGGCTCGCTAGAGGACCCAGAGGGCCCCGGAGCCCCCACCAGCCT
GCCACCGAGCGACTCCCCGTGCGCAAGGCCAGCAGCCACCGACGCGCCCTCCCGCCCCGGC
AGACTCGCAGGCAGGGTCCAAGCGTCCAGGTTTATTGACCCGGCTGGGTCTCACTCCTCCTT
CTCCTCCCCGTGGGTGATCACGTAGCTGAGCGCCTTG TAGTCCAGGTTGCCCGCCACATCGA
TGGAGGCGAACTGGAACATCTGGTCCACCTGCGGGCGGGGGCGAAAGGGCTCCTTGCGGGCT
CCGGGAGCGAATTACAAGCGCGCACCTGAAAA

FIGURE 169

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA50988

><subunit 1 of 1, 560 aa, 1 stop

><MW: 58427, pI: 6.86, NX(S/T): 2

MAGPTEAETGLAEPRALCAQRGHRTYARRWVFLLAISLLNCSNATLWLSFAPVADVIAEDLV
LSMEQINWLSLVYLVVSTPFGVAAIWILDSVGLRAATILGAWLNFAGSVLRMVPCMVVGTQN
PFAFLMGGQSLCALAQSLVIFSPAKLAALWFPEHQRTANMLATMSNPLGVLVANVLSPVLV
KKGEDIPMLGVYTIPAGVVCLLSTICLWESVPPTPPSAGAASSTSEKFLDGLKLQLMWNKA
YVILAVCLGGMIGISASFSALLEQILCASGHSSGFSGLCGALFITFGILGALALGPYVDRTK
HFTEATKIGLCLEFSLACVPFALVSQIQGQTLALAATCSLLGLFGFSVGPVAMELAVECSFPV
GEGAATGMIFVLGQAEGLIMLAMTALTVRREPSLSTCQQGEDPLDWTVSLLLMAGLCTFF
SCILAVFFHTPYRRLQAESGEPPSTRNAVGGADSGPGVDRGGAGRAGVLGPSTATPECTARG
ASLEDPRGPGSPHPACHRATPRAQGPAATDAPSRPGRLAGRVQASRFIDPAGSHSSFSSPWVIT

Important features:

Signal peptide:

amino acids 1-44

Transmembrane domains:

amino acids 61-79, 98-112, 126-146, 169-182, 201-215, 248-268,
280-300, 318-337, 341-357, 375-387, 420-441

N-glycosylation site.

amino acids 40-43 and 43-46

Glycosaminoglycan attachment site.

amino acids 468-471

[illegible]

GTCCACACATCTGCTCAACATGGGTGAGGTCCCTCTTAGACCAGCTCTTGTCATCATTTTGCTGAAGTGGACCAAC
TAGTTCCCCAGTAGGGGGTCTCCCTGGCAATTCTTGATCGGCGTTTGACATCTCAGATCGCTTCCAATGAAGA
TGGCCTTGCCCTTGGGGTCTGCTTGTTCATAATCATCTAATATATGGGACAAGGTTGTGCCGGCAGCTCTGGGGG
AAGGAGCACGGGGCTGATCAAGCCATCCAGGAACACATGGAGGACTTGTCCAGCCTTGAAAGAACTCTAGTGGTT
TCTGAATCTAGCCCATTTGGCGGTGAAGCATGATGCAACTTCTGCAACTTCTGCTGGGGCTTTTGGGGCCAGGTGG
CTACTTATTTCTTTTAGGGGATGTCAGGAGGTGACCACCTCTACGGTGAATACCAAGTGTAGAGGAAGTGGC
ATCTGGTACAGTGATCGGGAAGCTGTCCAGGAACCTGGGCCGGGAGGAGAGGCGGAGCAAGCTGGGGCCGCCTT
CCAGGTGTTGCAGCTGCCTCAGGCGCTCCCCATTAGGTGGACTCTGAGGAAGGCTTGTCTAGCACAGGCAGGCG
GCTGGATCGAGAGCAGCTGTGCCGACAGTGGGATCCCTGCCCTGGTTTCCCTTTGATGTGCTTGCCACAGGGGATTT
GGCTCTGATCCATGTGGAGATCCAAGTGTGGACATCAATGACCACCAGCCAGCGTTTCCCAAGGCGGAGCAGG
GCTGGAATCTCTGAGAGCGCCCTCTCTGCGAACCCGGATCCCCCTGGACAGAGCTCTTGACCCAGACACAGGCC
TAACACCCTGCACACCTACACTCTGTCTCCAGTGAACACTTTGCCCTGGATGTCAATTGTGGCCCTTGATGAGAC
CAAACATGCAGAACTCATAGTGGTGAAGGAGCTGGACAGGGAATCCATTCATTTTTTGATCTGGTGTAACTGC
CTATGACAATGGGAACCCCCCAAGTCAAGTACCAGCTTGGTCAAGGTCAACGTCCTTGGACTCCAATGACAATAG
CCCTGCGTTTTGCTGAGAGTTCACTGGCACCTGGAAATCCAAGAAGATGCTGCACCTGGTACGCTTCTCATAAACT
GACCGCCACAGACCCCTGACCAAGGCCCAATGGGGAGGTGGAGTTCTTCCCTCAGTAAGCACATGCCCTCCAGAGGT
GCTGGACACCTTGATATTGATGCCAAGACAGGCCAGGTCAATCTGCTGCCACCTCTAGACTATGAAAAGAACCC
TGCCCTACGAGTTGGATTTTACGCAAGGACCTGGGTCCCCAATCCTATCCAGCCCATGCAAGTTCTCAATCA
GGTTCTGGATGTCAATGACAACATCCCAAGCATCCACGTACATGGGCCCTCCAGCCATCAGTGTGTGAGAAGC
TCTTCCCAAGGACAGTTTTATTGCTCTTGTCTATGGCAGATGACTTGGATTAGGACACAATGGTTTGGTCCACTG
CTGGCTGAGCCCAAGAGCTGGGCCACTTCAGGCTGAAAAGAACTAATGGCAACACATACATGTTGCTAACCAATGC
CACACTGGACAGAGCAGTGGCCCAAAATATACCCCTCACTCTGTTAGCCCCAAGACCAAGGACTCCAGCCCTTATC
AGCCAAGAAACAGCTCAGACTTCAGATCAGTGCATCAACGACAATGCACCTGTGTTTGAGAAAAGCAGGTATGA
AGTCTCCACGCGGGAAAACAACTTACCCTCTCTTTCACCTCATTACCATCAAGGCTCATGATGACAGCTTGGGCAT
TAATGGAAAAGTCTCATACCGCATCCAGGACTCCCCAGTTGCTCACTTAGTAGCTATTGACTCCAACACAGGAGA
GGTCACTGCTCAGAGGTCACTGAACATATGAAGAGATGGCCGGCTTTGAGTTCCAGGTGATCGCAGAGGACAGCGG
GCAACCCATGCTTGCAATCCAGTGTCTCTGTGTGGGTGACGCTCTTGGATGCCAATGATAATGCCCCAGAGGTGGT
CCAGCTGTGCTCAGCAGTGGAAAAGCCAGCCTCTCCGTCTTGTGAATGCCCTCCACAGGCCACCTGCTGGTGGC
CATCGAGACTCCCAATGGCTTTGGGCCAGCGGGCACTGACACACTTCCATGGCCACTCAGACTCCCGCCCAT
CCTTTTGACAACCATTTGTGGCAAGAGATGCAGACTCGGGGGCAAATGGAGAGCCCCCTACAGCATCCGCAATGG
AAATGAAGCCCACCTCTTTCATCCTCAACCCCTCATACGGGGCAGCTGTTGCTCAATGTACCAATGCCAGCAGCCT
CATTGGGAGTGAGTGGGAGCTGGAGATAGTAGTAGAGGACCAGGGAAGCCCCCTTACAGACCCGAGCCCTGTT
GAGGCTCATGTTTGTCAACAGTGTGGACCACCTGAGGGACTCAGCCCGCAAGCCTGGGGCCCTTGAGCATGTCCGAT
GCTGACGTTGATCTGCTTGGCTGTACTGTTGGGCATCTTCGGTTGATCTCTGGCTTTGTTATGTTCCATCTGCCG
GACAGAAAAGAAGGACAACAGGGCCCTACAACTGTCTGGGAGGCGGATGCTCACTTACCGCCAGCAGCCCAAGAGGCC
CCAGAAACACATTCAGAAGGCAGACATCCACCTCGTGCCTGTGCTCAGGGGTGAGGCAGGTGAGCCTTGTGAAGT
CGGGCAGTCCCAAAAGATGTGGACAAGGAGGCGATGATGGAAGCAGGCTGGGACCCCTGCCTGCAGGCCCCCTT
CCACCTCACCCCGACCCCTGTACAGGACGCTGCGTAATCAAGGCAACCAGGGAGCACCGGCGGAGAGCCGAGAGGT
GCTGTCAAGACAGCGTCAACCTCTTTTCAACCATCCAGGCAGAGGAATGCCTCCCGGGAGAACCTGAACCTTCC
CGAGCCCCAGCCTGCCACAGGCCAGCCACGTTCCAGGCCCTGGAAGTTGCAAGGCACGCCACAGGAGGATGGC
TGGAGACCAGGGCAGTGAGGAAGCCCCACAGAGGCCACCAGCCTCCTCTGCAACCTGAGACCGGCAGGACGATCT
CAATGGCAAAGTGTCCCTTGAGAAAGAATCAGGGCCCCGTGAGATCCTGCGGAGCCTGGTCCGGCTGTCTGTGGC
TGCCCTTCGCCGAGCGGAACCCCGTGGAGGAGCTCACTGTGGATTCTCCTCCTGTTTACGAAATCTCCAGCTGCT
GTCTTGTGCTGATCAGGGCCAAATCCAGGCCAAACAAACACCCGAGGAAATAAGTACTTGGCCAAGCCAGGAGG
CAGCAGTGTGCAATCCAGGACAGATGGCCCAAGTGCAAGGCTGGAGGCGAGACGACCCAGAACAGGAGGA
AGGCCTTTGGATCCTGAAGAGGACCTCTCTGTGAAGCAACTGCTAGAAGAAGAGCTGTCAAGTCTGCTGGACCC
CAGCACAGGTCTGGCCCTGGACCGGCTGAGCGCCCCCTGACCCGGCCTGGATGGCGAGACTCTCTTTGCCCTCAC
CACCACCTACCGTGACAATGTGATCTCCCCGGATGCTGCAGCCACGAGGAGCCGAGGACCTTCCAGACGTTCCG
CAAGGCAGAGGCACCCAGAGCTGAGCCCAACAGGCACGAGGCTGGCCAGCACCTTTGTCTCGGAGATGAGCTCACT
GCTGGAGATGCTGCTGGAACAGCGCTCCAGCATGCCCGTGGAGGCCGCTCCGAGGCGCTGCGGCGGCTCTCGGT
CTGCGGGAGGACCTCAGTTTAGACTTGGCCACCAGTGACGCTCAGGCTGAAAGTGCAAGGGGACCCAGGTGG
AAAGACGGGGACTGAGGGTAAAGAGCAGGCGACAGCAGCAGCAGGATGCCTGTGAACATACCTCAGAGCCT
CTGGATCCAAGAACCAGGGGCTGAGGATCTGTGGACAAGAGCTGGTTTCTAAAACTTTGTAACCTACTAGCTAG
CGGCGGCTTGAGAACTTTAGGGTGAAGTGTCTACCCCCACAGAGGAGGCAAGAGCCCCAGGACTAACAGCTGAC
TGACCAAGACAGCCCTTGTAAGCAGCTCTGAGTCTTTTGGAGGACAGGGACGGTTTGTGGCTGAGATAAGTGT
TCTTGGCAAAACATATGTGGAGCACAAGGGTCAGTCTCTGGCAGAACAGATGCCACGGAGTATCACAGGCAGG
AAAGGTGGCCCTTCTTGGGTAGCAGGAGTCAGGGGCTGTACCCTGGGGGTGCCAGGAATGCTCTCTGACCTAT
CAATAAAGGAAAAGCAGTAAAAAAAAAAAAAAAAAAAAA

FIGURE 171

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48331

<subunit 1 of 1, 1184 aa, 1 stop

<MW: 129022, pI: 5.20, NX(S/T): 5

MMQLLQLLLGLLGPGGYLFLLGDCQEVTTLTVKYQVSEEVPSGTVIGKLSQELGREERRRQA
GAAFQVLQLPQALPIQVDSEEGLLSTGRRLDREQLCRQWDPCLVSFDVLATGDLALIHVEIQ
VLDINDHQPRFPKGEQELEISESASLRTRIPLDRALDPTGPNLTHTYTLSPSEHFALDVIV
GPDETKHAELIVVKELDREIHSFFDLVLTAYDNGNPPKSGTSLVKVNVLDSDNDNSPAFAESS
LALEIQEDAAPGTLLIKLTATDPDQGPNGEVEFFLSKHMPPEVLDTFSIDAKTGQVILRRPL
DYEKNPAYEVDVQARDLGPNPPIPAHCKVLIKVLVDNDNIPSIHVTWASQPSLVSEALPKDSF
IALVMADDLDSGHNGLVHCWLSQELGHFRLKRTNGNTYMLLTNATLDREQWPKYTLTLLAQD
QGLQPLSAKKQLSIQISDINDNAPVFEKSRYEVSTRENNLPSLHLITIKAHDADLGINGKVS
YRIQDSPVAHLVAIDSNTGEVTAQRSLNYYEEMAGFEFQVIAEDSGQPMLASSVSVVWSLLDA
NDNAPEVVQPVLSDGKASLSVLVNASTGHLLVPIETPNGLGPAGTDTTPPLATHSSRPFLTT
IVARDADSGANGEPLYSIRNGNEAHLFILNPHTGQLFVNVTNASSLIGSEWELEIVVEDQGS
PPLQTRALLRVMFVTSVDHLRDSARKPGALSMSMLTVICLAVLLGIFGLILALFMSICRTEK
KDNRAYNCREAESTYRQQPKRPQKHIQKADIHLVPVLRGQAGEPCEVGQSHKDVDKEAMMEA
GWDPCLOAPFHLTPTLYRTLNRQGNQGAPAESREVLQDTVNLLFNHPRQRNASRENLNLP
QPATGQPRSRPLKVAGSPTGRLAGDQGSEEAPQRPPASSATLRRQRHLNGKVSPEKESGPRQ
ILRSLVRLSVAAFAERNPVEELTVDSPPVQQISQLLSLLHQGFQPKPNHRGNKYLAKEGGS
RSAIPDTDGPSARAGGQTDPEQEEGPLDPEEDLSVKQLLEEELSSLLDPSTGLALDRLSAPD
PAWMARLSLPLTTNYRDNVISPDAAATEEPRTFQTFGKAEAPELSPTGTRLASTFVSEMSSL
LEMLLEQRSSMPVEAASEALRRLSVCGRTLSDLATSAASGMKVQGDPPGKTGTGEGKSRGSS
SSSRCL

Important features:

Signal peptide:

amino acids 1-13

Transmembrane domain:

amino acids 719-739

N-glycosylation site.

amino acids 415-418, 582-585, 659-662, 662-665 and 857-860

Cadherins extracellular repeated domain signature.

amino acids 123-133, 232-242, 340-350, 448-458 and 553-563

FIGURE 172

CGGACGCGTGGGCGGACGCGTGGGGGAGAGCCGCAGTCCCGGCTGCAGCACCTGGGAGAAGG
CAGACCGTGTGAGGGGGCCTGTGGCCCCAGCGTGCTGTGGCCTCGGGGAGTGGGAAGTGGAG
GCAGGAGCCTTCCTTACACTTCGCCATGAGTTTCCTCATCGACTCCAGCATCATGATTACCT
CCCAGATACTATTTTTTGGATTTGGGTGGCTTTTCTTCATGCGCCAATTGTTTAAAGACTAT
GAGATACGTCAGTATGTTGTACAGGTGATCTTCTCCGTGACGTTTGCATTTTCTTGACCCAT
GTTTGAGCTCATCATCTTTGAAATCTTAGGAGTATTGAATAGCAGCTCCCGTTATTTTCACT
GGAAAATGAACCTGTGTGTAATTCTGCTGATCCTGGTTTTTCATGGTGCCTTTTTACATTGGC
TATTTTATTGTGAGCAATATCCGACTACTGCATAAACAACGACTGCTTTTTTCTGTCTCTT
ATGGCTGACCTTTATGTATTTCTTCTGGAACTAGGAGATCCCTTTCCCATTCTCAGCCCAA
AACATGGGATCTTATCCATAGAACAGCTCATCAGCCGGGTTGGTGTGATTGGAGTGACTCTC
ATGGCTCTTCTTTCTGGATTTGGTGTCTGTCAACTGCCCATACACTTACATGTCTTACTTCT
CAGGAATGTGACTGACACGGATATTCTAGCCCTGGAACGGCGACTGCTGCAAACCATGGATA
TGATCATAAGCAAAAAGAAAAGGATGGCAATGGCACGGAGAACAAATGTTCCAGAAGGGGGAA
GTGCATAACAAACCATCAGGTTTCTGGGGAATGATAAAAAGTGTTACCACTTCAGCATCAGG
AAGTGAAAATCTTACTCTTATTCAACAGGAAGTGGATGCTTTGGAAGAATTAAGCAGGCAGC
TTTTTCTGGAAACAGCTGATCTATATGCTACCAAGGAGAGAATAGAATACTCCAAAACCTTC
AAGGGGAAATATTTTAATTTTCTTGGTTACTTTTTCTCTATTTACTGTGTTTGGAAAATTTT
CATGGCTACCATCAATATTGTTTTTGATCGAGTTGGGAAAACGGATCCTGTCACAAGAGGCA
TTGAGATCACTGTGAATTATCTGGGAATCCAATTTGATGTGAAGTTTGGTCCCAACACATT
TCCTTCATTCTTGTTGGAATAATCATCGTCACATCCATCAGAGGATTGCTGATCACTCTTAC
CAAGTTCTTTTATGCCATCTCTAGCAGTAAGTCCTCCAATGTCATTGTCCTGCTATTAGCAC
AGATAATGGGCATGTACTTTGTCTCCTCTGTGCTGCTGATCCGAATGAGTATGCCTTTAGAA
TACCGCACCATAATCACTGAAGTCCTTGGAGAACTGCAGTTCAACTTCTATCACCGTTGGTT
TGATGTGATCTTCCTGGTCAGCGCTCTCTCTAGCATACTCTTCCTCTATTTGGCTCACAAAC
AGGCACCAGAGAAGCAAATGGCACCTTGAACTTAAGCCTACTACAGACTGTTAGAGGCCAGT
GGTTTCAAATTTAGATATAAGAGGGGGGAAAAATGGAACCAGGGCCTGACATTTTATAAAC
AAACAAAATGCTATGGTAGCATTTTTTCACCTTCATAGCATACTCCTTCCCCGTCAGGTGATA
CTATGACCATGAGTAGCATCAGCCAGAACATGAGAGGGAGAACTAACTCAAGACAATACTCA
GCAGAGAGCATCCCGTGTGGATATGAGGCTGGTGTAGAGGCGGAGAGGAGCCAAGAACTAA
AGGTGAAAAATACACTGGAACCTCTGGGGCAAGACATGTCTATGGTAGCTGAGCCAAACACGT
AGGATTTCCGTTTTAAGGTTACATGGAAAAGGTTATAGCTTTGCCTTGAGATTGACTCATT
AAAATCAGAGACTGTAACAAAAAAAAAAAAAAAAAAAAAGGGCGGCCGCGACTCTAGAGTCG
ACCTGCAGAAGCTTGGCCGCCATGGCCCAACTTGTTTATTGCAGCTTATAATG

FIGURE 173

MSFLIDSSIMITSQILFFGFGWLFFMRQLFKDYEIRQYVVQVIFSVTFASFCTMFELIIFEI
LGVLNSSSRYPFHWMNLCVILLILVFMVPFYIGYFIVSNIRLLHKQRLLFSCLLWLTFMYFF
WKLGDPPILSPKHGILSIEQLISRVGVIGVTLMALLSGFGAVNCPYTYMSYFLRNVTDTDI
LALERRLLQTMDMIISKKKRMAMARRTMFQKGEVHNKPSGFWGMIKSVTTSASGSENLTLIQ
QEVDALEELSRQLFLETADLYATKERIEYSKTFKGKYFNFLGYFFSIYCVWKIFMATINIVF
DRVGKTDVPVTRGIEITVNYLGIQFDVKFWSQHISFILVGIIIVTSIRGLLITLTKFFYAIS
SKSSNVIVLLLAQIMGMYFVSSVLLIRMSMPLEYRTIITEVLGELQFNFYHRWFDVIFLVSA
LSSILFLYLAHKQAPEKQMAP

Important features:

Signal peptide:

amino acids 1-23

Potential transmembrane domains:

amino acids 37-55, 81-102, 150-168, 288-311, 338-356, 375-398,
425-444

N-glycosylation sites.

amino acids 67-70, 180-183 and 243-246

Eukaryotic cobalamin-binding proteins

amino acids 151-160

FIGURE 174

CATGGGAAGTGGAGCCGGAGCCTTCCTTACACTCGCCATGAGTTTCCTCATCGACTCCAGCA
TCATGATTACCTCCCNGANACTATTTTTTTGGATTTGGGTGGCTTTTCTTCNGCGCCAATGTT
TAAAGACTATGAGATACGTCAGTATGTTGTACNNGTGATCTTCTCCGTGACGTTTGCCATTT
CTTGCAACCATGTTTGAGCTCATCATCTTTGAAATCTTNGGAGTATTGAATAGCAGCTCCCGT
TATTTTCACTGGAAAATGAACCTGTGTGTAATTCTGCTGATCCTGGTTNTCATGGTGCCTTT
TTACATTGGCTATTTTATTGTGAGCAATATCCGACTACTGCATAAACGACTGCTTTTTT
CCTGTCTCTTATGGCTGACCTTTATGTATTTCCAG

FIGURE 175

GTGTTGCCCTTGGGGAGGGGAAGGGGAGCCNGGCCCTTTCCTAAAATTTGGCCAAGGGTTTC
TTTNTTGAATTCCGGGTNNNGNATACCTTCCCAGAAAATATTTTTTGGATTTGGGGTAGNTT
TTTTTCATGCGCCAATTGTTTAAAGACTATGAGATACGTCAGTATGTTGTACAGGTGATNTT
NTCCGTGACGTTTGCATTTTCTTGCACCATGTTTGAGCTCATCATNTTTGAAATNTTAGGAG
TATTGAATAGCAGCTCCCGTTATTTTCACTGGAAAATGAACCTGTGTGTAATTCTGCTGATC
CTGGTTTTTCATGGTGCCTTTTTTACATTGGCTATTTTATTGTGAGCAATATCCGACTACTGCA
TAAACAACGACTGCTTTTTTCTGTCTNTTATGGCTGACCTTTATGTATTTNTTNTGGAAAN
TAGGAGATCCCTTTCCCATTCCTC

1000967 102401

FIGURE 176

CTCGCGCAGGGATCGTCCCATGCGCGGGGCTCGGAGCCGCGACCCCTTGGGGGGCCTCCGGGATTTGCTACCTTTT
TGGCTCCCTGCTCGTGAAGTCTCTTCTCACGGGCTGTGCGCTTCAATCTGGACGTGATGGGTGCTTTGCGCAA
GGAGGGCGAGCCAGGCAGCCTCTTCGGCTTCTCTGTGGCCCTGCACCGGCAGTTGCAGCCCCGACCCAGAGCTG
GCTGCTGGTGGGTGCTCCCCAGGCCCTGGCTCTTCTGGGCGAGCAGGCGAATCGCACTGGAGGCCTCTTCGCTTG
CCCGTTGAGCCTGGAGGAGACTGACTGCTACAGAGTGGACATCGACCAGGGAGCTGATATGCAAAAGGAAAGCAA
GGAGAACCAGTGGTTGGGAGTCAGTGTTCGGAGCCAGGGGCCCTGGGGGCAAGATTGTTACCTGTGCACACCCGATA
TGAGGCAAGGCAGCGAGTGGACAGATCCTGGAGACGCGGATATGATTGGTTCGCTGCTTTGTGCTCAGCCAGGA
CCTGGCCATCCGGGATGAGTTGGATGGTGGGGAATGGAAGTTCTGTGAGGGACGCCCCAAGGCCATGAACAATT
TGGGTTCTGCCAGCAGGGCACAGCTGCCGCCCTTCTCCCCTGATAGCCACTACCTCCTCTTTGGGGCCCCAGGAAC
CTATAATTGGAAGGGCACGGCCAGGGTGGAGCTCTGTGCACAGGGCTCAGCGGACCTGGCACACCTGGACGACGG
TCCCTACGAGGCGGGGGGAGAGAAGGAGCAGGACCCCCGCTCATCCCGGTCCCTGCCAACAGCTACTTTGGCTT
CTCTATTGACTCGGGGAAAGGTCTGGTGCCTGCAGAAAGAGCTGAGCTTTGTGGCTGGAGCCCCCGCGCCAAACCA
CAAGGGTGTCTGGTTCATCCTGCGCAAGGACAGCGCCAGTGCCTGGTGGCCGAGGTTATGCTGTCTGGGGAGCG
CCTGACCTCCGGCTTTGGCTACTCACTGGCTGTGGCTGACCTCAACAGTGTGCTGGCTGGCCAGACCTGATAGTGGG
TGCCCCCTACTTTTGGAGCGCAAGAAGAGCTGGGGGGTGTGTGTATGTGTACTTGAACAGGGGGGTCACTG
GGCTGGGATCTCCCCTCTCCGGCTCTGCGGCTCCCCTGACTCCATGTTCCGGATCAGCCTGGCTGTCTGGGGGA
CCTCAACCAAGATGGCTTTCCAGATATTGAGTGGGTGCCCCCTTTGATGGTGTGGGAAAGTCTTCATCTACCA
TGGGAGCAGCCTGGGGGTTGTGCGCAAAACCTTACAGGTGCTGGAGGGCGAGGCTGTGGGCATCAAGAGCTTCGG
CTACTCCCTGTGAGCAGCTTGGATATGGATGGGAACCAATACCTGACCTGCTGGTGGGCTCCCTGGCTGACAC
CGCAGTGTCTTTCAGGGCCAGACCCATCCTCCATGTCTCCATGAGGTCTCTATTGCTCCACGAAGCATCGACCT
GGAGCAGCCCCAAGTGTGCTGGCGGCCACTCGGTCTGTGTGGACCTAAGGGTCTGTTTCAGCTACATTCAGTCCC
CAGCAGCTATAGCCCTACTGTGGCCCTGGACTATGTGTTAGATGCGGACACAGACCGGAGGCTCCGGGGCCAGGT
TCCCCGTGTGACGTTCTGAGCCGTAACCTGGAAGAACCAGCACCAGGCCTCGGGCACCGTGTGGCTGAAGCA
CCAGCATGACCGAGTCTGTGGAGACGCCATGTTCCAGCTCCAGGAAAATGTCAAAGACAAGCTTCGGGCCATTTGT
AGTGACCTTGTCTACAGTCTCCAGACCCCTCGGCTCCGGCGACAGGCTCCTGGCCAGGGGTGCTCCAGTGGC
CCCCATCCTCAATGCCACCAGCCAGCACCAGCGGGCAGAGATCCACTTCTGGAAGCAAGGCTGTGGTGAGA
CAAGATCTGCCAGAGCAATCTGCAGCTGGTCCACGCCCGCTTCTGTACCCGGGTGAGCGACACGGAATTCGAACC
TCTGCCCATGGATGTGGATGGAACAACAGCCCTGTTTGCAGTGTGGGCGAGCCAGTCAATTGGCCTGGAGCTGAT
GGTCAACCACTGCCATCGGACCCAGCCAGCCCGAGGCTGATGGGGATGATGCCCATGAAGCCAGCTCCTGGT
CATGCTTCTGACTCACTGCACTACTCAGGGGTCCGGGCCCTGGACCCCTGCGGAGAAGCCACTCTGCCTGTCCAA
TGAGAATGCCTCCCATGTTGAGTGTGAGCTGGGGAACCCATGAAGAGAGGTGCCAGGTCACTTCTACCTCAT
CCTTAGCACCTCCGGGATCAGCATTGAGACCAGGAACCTGGAGGTAGAGCTGCTGTTGGCCACGATCAGTGAGCA
GGAGCTGCATCCAGTCTCTGCACGAGCCCGTGTCTTTCATTGAGCTGCCACTGTCCATTGCAGGAATGGCCATTCC
CCAGCAACTCTTCTTCTCTGGTGTGGTGGGGGCGAGAGGCCATGCAGTCTGAGCGGGATGTGGGCAGCAAGGT
CAAGTATGAGGTACGGTTTCCAACCAAGGCCAGTGCCTCAGAACCCTGGGCTCTGCCTTCTCAACATCATGTG
GCCTCATGAGATTGCCAATGGGAAGTGGTTGCTGTACCCAATGCAGGTTGAGCTGGAGGGCGGGCAGGGGCCCTGG
GCAGAAAGGGCTTTGCTCTCCAGGCCCAACATCCTCCACCTGGATGTGGACAGTAGGGATAGGAGGCGGCGGGA
GCTGGAGCCACCTGAGCAGCAGGAGCCTGGTGAGCGGCAGGAGCCAGCATGTCTGGTGGCCAGTGTCTCTGTC
TGAGAAGAAGAAAAACATCACCTGGACTGCGCCCGGGGCACGGCCAACTGTGTGGTGTTCAGCTGCCCACTCTA
CAGCTTTGACCGCGCGGCTGTGCTGCATGTCTGGGGCCGTCTCTGGAACAGCACCTTTCTGGAGGAGTACTCAGC
TGTGAAGTCCCTGGAAGTGATTGTCCGGGCCAACATCACAGTGAAGTCTCCATAAAGAACTTGATGCTCCGAGA
TGCTTCCACAGTGATCCAGTGATGGTATACTTGGACCCCATGGCTGTGGTGGCAGAAGGAGTGCCCTGGTGGGT
CATCTCCTGGCTGTACTGGCTGGGCTGTGGTGTAGCACTGTGGTGTGCTCCTGTGGAAGATGGGATTTCTT
CAAACGGGCGAAGCACCCCGAGGCCACCGTGCCCCAGTACCATGCGGTGAAGATTCTCGGGAAGACCCGACAGCA
GTTCAAGGAGGAGAAGACGGGCACCATCCTGAGGAACAACCTGGGGCAGCCCCCGGCGGGAGGGCCCCGATGCACA
CCCCATCCTGGCTGCTGACGGGCATCCCGAGCTGGGCCCCGATGGGCATCCAGGGCCAGGCACCGCTAGGTTCC
CATGTCCCAGCCTGGCCTGTGGCTGCCCTCCATCCCTTCCCCAGAGATGGCTCCTTGGGATGAAGAGGGTAGAGT
GGGCTGCTGGTGTGCGATCAAGATTTGGCAGGATCGGCTTCTCAGGGGCACAGACCTCTCCCAACCAAGAAC
TCCTCCCACCAACTTCCCCTTAGAGTGTGTGAGATGAGAGTGGGTAAATCAGGGACAGGGCCATGGGGTAGGG
TGAGAAGGGCAGGGGTGTCTGATGCAAGGTGGGGAGAAGGGATCCTAATCCCTTCTCTCCATTACCCCTGT
GTAACAGGACCCCAAGGACCTGCCTCCCCGGAAGTGCTTAACTAGAGGGTGGGGAGGAGGTTGTGTCACTGA
CTCAGGCTGCTCCTTCTCTAGTTTCCCCTCTCATCTGACCTTAGTTTGTGCTGCCATCAGTCTAGTGGTTTCGTGGT
TTCGTCTATTTATTAAAAAATATTTGAGAACAAAAA

FIGURE 177

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA55737

><subunit 1 of 1, 1141 aa, 1 stop

><MW: 124671, pI: 5.82, NX(S/T): 5

MAGARSRDPWGASGICYLFGSLLVELLFSRAVAFNLDVMGALRKEGEPGSLFGFSVALHRQL
QPRPQSWLLVGAPQALALPGQQANRTGGLFACPLSLEETDCYRVDIDQGADMQKESKENQWL
GVSVRSQGPQGGKIVTCAHRYEARQRVDQILETRDMIGRCFVLSQDLAIRDELDDGGGEWKFC
RPQGHEQFGFCQQGTAAAFSPDSHYLLFGAPGTYNWKGRTARVELCAQGSADLAHLDDGPYEA
GGEKEQDPRLIPVPANSYFGFSIDSGKGLVRAEELS FVAGAPRANHKGAVVILRKDSASRLV
PEVMLSGERLTSGFGYSLAVADLNSDGPDLIVGAPYFFERQEELGGAVVYVLNQGGHWAGI
SPLRLCGSPDSMFGISLAVLGDNLQDGFDPDIAVGAPFDGDGKVFYIHGSSLGVVAKPSQVLE
GEAVGIKSFGYSLSGSLDMDGNQYPDLLVGSLADTAVLFRARPILHVSHEVSIAPRSIDLEQ
PNCAGGHSVCVDLRVCFSYIAVPSSYSPTVALDYVLDADTDRLRGQVPRVTFLSRNLEEPK
HQASGTVWLKHQHDRVCGDAMFQLQENVKDKLRAIVVTLSSYSLQTPRLRRQAPGQGLPPVAP
ILNAHQSTQRAEIHFLKQGCGEDKICQSNLQLVHARFCTRVSDTEFQPLPMDVDGTTALFA
LSGQPVIGLELMVTNLPSPDPAQPQADGDDAHEAQLLVMLPDSLHYSGVRALDPAEKPLCLSN
ENASHVECELG NPMKRGAVTFYILISTSGISIIETTELEVELLATISEQELHPVSARARVF
IELPLSIAGMAIPQQLFFSGVVRGERAMQSERDVGSKVKYEVTVSNQGQSLRTLGS AFLNIM
WPHEIANGKWLLYPMQVELEGGQGPQKGLCSPRPNILHLDVDSRDRRRRELEPPEQQEPGE
RQEPSMSWWPVSSAEKKKNITLDCARGTANCVVFSCLYSDRAAVLHVWGRLWNSTFLEEY
SAVKSLEIVIVRANITVKSSIKNLMLRDASTVIPVMVYLDPMVVAEGVPWWVILLAVLAGLL
VLALLVLLLWKMGFFKRAKHPEATVPQYHAVKIPREDRQQFKEEKTGTILRNNWGSPRREGP
DAHPIAADGHPGPDGHPGPGTA

Important features:

Signal peptide:

amino acids 1-33

Transmembrane domain:

amino acids 1040-1062

N-glycosylation sites.

amino acids 86-89, 746-749, 949-952, 985-988 and 1005-1008

Integrins alpha chain proteins.

amino acids 1064-1071, 384-408, 1041-1071, 317-346, 443-465, 385-407, 215-224, 634-647, 85-99, 322-346, 470-479, 442-466, 379-408 and 1031-1047

FIGURE 178

CGCGCCGGGCGCAGGGAGCTGAGTGGACGGCTCGAGACGGCGGCGCGTGCAGCAGCTCCAGA
AAGCAGCGAGTTGGCAGAGCAGGGCTGCATTTCCAGCAGGAGCTGCGAGCACAGTGCTGGCT
CACAACAAGATGCTCAAGGTGTCAGCCGTACTGTGTGTGTGTGCAGCCGCTTGGTGCAGTCA
GTCTCTCGCAGCTGCCGCGGCGGTGGCTGCAGCCGGGGGGCGGTCCGACGGCGGTAATTTTC
TGGATGATAAAACAATGGCTCACCACAATCTCTCAGTATGACAAGGAAGTCGGACAGTGGAAC
AAATTCGAGACGAAGTAGAGGATGATTATTTCCGCACTTGGAGTCCAGGAAAACCCCTTCGA
TCAGGCTTTTAGATCCAGCTAAGGATCCATGCTTAAAGATGAAATGTAGTCGCCATAAAGTAT
GCATTGCTCAAGATTCTCAGACTGCAGTCTGCATTAGTCACCGGAGGCTTACACACAGGATG
AAAGAAGCAGGAGTAGACCATAGGCAGTGGAGGGGTCCCATATTATCCACCTGCAAGCAGTG
CCCAGTGGTCTATCCCAGCCCTGTTTGTGGTTCAGATGGTCATACCTACTCTTTTCAGTGCA
AACTAGAATATCAGGCATGTGTCTTAGGAAAACAGATCTCAGTCAAATGTGAAGGACATTGC
CCATGTCCTTCAGATAAGCCCACCAGTACAAGCAGAAATGTTAAGAGAGCATGCAGTGACCT
GGAGTTCAGGGAAGTGGCAAACAGATTGCGGGACTGGTTCAAGGCCCTTCATGAAAGTGGAA
GTCAAAACAAGAAGACAAAAACATTGCTGAGGCCTGAGAGAAGCAGATTTCGATACCAGCATC
TTGCCAATTTGCAAGGACTCACTTGGCTGGATGTTTAAACAGACTTGATACAACTATGACCT
GCTATTGGACCAGTCAGAGCTCAGAAGCATTTACCTTGATAAGAATGAACAGTGTACCAAGG
CATTCTTCAATTCTTGTGACACATACAAGGACAGTTTAATATCTAATAATGAGTGGTGCTAC
TGCTTCCAGAGACAGCAAGACCCACCTTGCCAGACTGAGCTCAGCAATATTCAGAAGCGGCA
AGGGGTAAAGAAGCTCCTAGGACAGTATATCCCCCTGTGTGATGAAGATGGTTACTACAAGC
CAACACAATGTCATGGCAGTGTGGACAGTGTGGTGTGTTGACAGATATGGAAATGAAGTC
ATGGGATCCAGAATAAAATGGTGTTCAGATTGTGCTATAGATTTTGAGATCTCCGGAGATTT
TGCTAGTGGCGATTTTCATGAATGGACTGATGATGAGGATGATGAAGACGATATTATGAATG
ATGAAGATGAAATTGAAGATGATGATGAAGATGAAGGGGATGATGATGATGGTGGTGTGAC
CATGATGTATACATTTGATTGATGACAGTTGAAATCAATAAAATCTACATTTCTAATATTTA
CAAAAATGATAGCCTATTTAAAATTATCTTCTTCCCCAATAACAAAATGATTCTAAACCTCA
CATATATTTTGTATAATTATTTGAAAAATTGCAGCTAAAGTTATAGAACTTTATGTTTAAAT
AAGAATCATTTGCTTTGAGTTTTTATATTCCTTACACAAAAGAAAATACATATGCAGTCTA
GTCAGACAAAATAAAGTTTTGAAGTGCTACTATAATAAATTTTTTCACGAGAACAACTTTGT
AAATCTTCCATAAGCAAAATGACAGCTAGTGCTTGGGATCGTACATGTTAATTTTTTTGAAAG
ATAATTCTAAGTGAAATTTAAAATAAATAAATTTTTAATGACCTGGGTCTTAAGGATTTAGG
AAAAATATGCATGCTTTAATTGCATTTCCAAAGTAGCATCTTGCTAGACCTAGATGAGTCAG
GATAACAGAGAGATACCACATGACTCCAAAAA

FIGURE 179

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA49829

><subunit 1 of 1, 436 aa, 1 stop

><MW: 49429, pI: 4.80, NX(S/T): 0

MLKVSAVLCVCAAAWCSQSLAAAAVAAAGGRSDGGNFLDDKQWLTTISQYDKEVGQWNKFR
DEVEDDYFRTWSPGKPFQALDPAKDPCLKMKCSRHKVCIAQDSQTAVCISHRRLTHRMKEA
GVDHRQWRGPILSTCKQCPVVYPSPVCGSDGHTYSFQCKLEYQACVLGKQISVKCEGHCP
SDKPTSTSRNVKRACSDLEFREVANRLRDWFKALHESGSQNKKTLLRPERSRFDTSILPI
CKDSLGMFNRDLDTNYDLLLDQSELRSIYLDKNEQCTKAFFNSCDTYKDSLISNNEWCYCFQ
RQQDPPCQTELSNIQKRQGVKLLGQYIPLCDEDEGYKPTQCHGSVGQCWCVDTRYGNEVMGS
RINGVADCAIDFEISGDFASGDFHEWTDDEDEDDIMNDEDEIEDDDEDEGDDDDGGDDHDVYI

Important features:

Signal peptide:

amino acids 1-16

Leucine zipper pattern.

amino acids 246-267

N-myristoylation sites.

amino acids 357-362, 371-376 and 376-381

Thyroglobulin type-1 repeat proteins

amino acids 353-365 and 339-352

FIGURE 180

CAGACTCCAGATTTCCCTGTCAACCACGAGGAGTCCAGAGAGGAAACGCGGAGCGGAGACAACAGTACCTGACGC
 CTCTTTTCAGCCCGGGATCGCCCCAGCAGGGATGGGCGACAAGATCTGGCTGCCCTTCCCCGTGCTCCTTCTGGCC
 GCTCTGCCTCCGGTGCTGCTGCCTGGGGCGGCCGGCTTCACACCTTCCCTCGATAGCGACTTCACCTTTACCCCTT
 CCGCCCGGCCAGAAGGAGTGCTTCTACCAGCCCATGCCCTGAAGGCCTCGCTGGAGATCGAGTACCAAGTTTGA
 GATGGAGCAGGATTAGATATTGATTTCCATCTTGCCCTCTCCAGAAGGCAAAACCTTAGTTTTTGAACAAAGAAAA
 TCAGATGGAGTTCACACTGTAGAGACTGAAGTTGGTGATTACATGTTCTGCTTTGACAATACATTTCAGCACCATT
 TCTGAGAAGGTGATTTTCTTTGAATTAATCCTGGATAATATGGGAGAACAGGCACAAGAACAAAGATTGGAAG
 AAATATATTACTGGCACAGATATATTGGATATGAAACTGGAAGACATCCTGGAATCCATCAACAGCATCAAGTCC
 AGACTAAGCAAAAGTGGGCACATACAAATCTGCTTAGAGCATTTGAAGCTCGTGATCGAAACATACAAGAAAGC
 AACTTTGATAGAGTCAATTTCTGGTCTATGGTTAATTTAGTGGTCAATGGTGGTGGTGTGAGCCATTCAAGTTTAT
 ATGCTGAAGAGTCTGTTTGAAGATAAGAGGAAAAGTAGAATCTAAACTCCAACTAGAGTACGTAACATTGAAA
 AATGAGGCATAAAATGCAATAAACTGTTACAGTCAAGACCATTAAATGGTCTTCTCCAAAATATTTTGAGATATA
 AAAGTAGGAAAACAGGTATAATTTAATGTGAAAATTAAGTCTTCACTTCTGCTGCAAGTAATCCTGCTGATCCAG
 TTGTACTTAAAGTGTGTAACAGGAATATTTGTCAGAAATATAGGTTTAACTGAATGAAGCCATATTAATAACTGCAT
 TTTCTAACTTTGAAAAATTTTGCAAAATGCTTAGGTGATTTAAATAAAATGAGTATTGGGCCTAATTGCAACACC
 AGTCTGTTTTTAACAGGTTCTATTACCCAGAACTTTTTTGTAATGCGGCAGTTACAAATTAAGTGTGGAAGTTT
 TCAGTTTTAAGTTATAAATCACCTGAGAATTACCTAATGATGGATTGAATAAATCTTTAGACTACAAAAGCCCAA
 CTTTTCTCTATTTACATATGCATCTCTCCTATAATGTAATAGATAATAGCTTTGAAATACAATTAGGTTTTTG
 AGATTTTTATAACCAAATACATTTCAAGTGAACATATTAGCAGAAAGCATTAGTCTTTGTACTTTGCTTACATTC
 CCAAAGCTGACATTTTACGATTCTTAAAAACACAAAGTTACACTTACTAAAATTAGGACATGTTTTCTCTTTG
 AAATGAAGAATATAGTTTTAAAAGCTTCCCTCCATAGGGACACATTTTCTTAACCCCTTAAGTAAAGTGTAGGA
 TTTTAAAATTAAGTGTGAGGTAATAAAGTTTATTTTTAATAGTATCTGTCAAGTTAATATCTGTCAACAGTTAA
 TAATCATGTTATGTTAATTTTAACATGATTGCTGACTTGGATAAATCATTATTACCAGCAGTTATGAAGGAAATA
 TTGCTAAAATGATCTGGGCCTACCATAAAATAAATATCTCCTTTTCTGAGCTCTAAGAATTATCAGAAAACAGGAA
 AGAATTTAGAAAACTTGAGAAAACCTAATCCAAAATAAAATTCACTTAAGTAGAACTATAAATAAATATCTAGA
 ATCTGACTGGCTCATCTGACATCCTACTCATAACATAAATCAAAGGAGATGATTAATTTCCAGTTAGCTGGAAG
 AAACCTTGGCTGTAGGTTTTTATTTTCTACAAGAATTTCTGGTTTGAATTATTTTGTAAGCAGGTACATTTTATA
 AAATGTAAGCCCTACTGTAAGGTTTAGCACTGGGTGTACATATTTATTAATAAATTTTATTATAACAACCTTTTAT
 TAAAATGGCCTTTCTGAACACTTTATTTATTTGATGTTGAAGTAAGGATTAGAAACATAGACTCCCAAGTTTTAAA
 CACCTAAATGTGAATAACCCATATATACACAAAGTTTCTGCCATCTAGCTTTTTGAGTCTATGGGGGTCTTAC
 TCAAGTACTAGTAATTTAACTTCATCATGAATGAACATAAATTTTAAAGTTATGCCATTTATAACGTTGTTTAT
 GACTACATTGTGAGTTAGAAAACAACTTAAATTTGGGGTATAGAACCCTCAACAGGTTAGTAATGCTGGAATT
 CTTGATGAGCAATAATGATAACCAGAGAGTGATTTTCATTTACACTCATAGTAGTATAAAAAGAGATACATTTCCC
 TCTTAGGCCCCCTGGGAGAAGAGCAGCTTAGATTTCCCTACTGGCAAGGTTTTTAAAAATGAGGTAAATGCCGTAT
 ATGATCAATTACCTTAATTGGCCAAGAAAATGCTTCAGGTGTCTAGGGGTATCCTCTGCAACACTTGCAGAACAA
 AGGTCAATAAGATCCTTGCCATGAATACCCCTCCCTTTTGCGCTGTTAAATTTGCAATGAGAAGCAAATTTACA
 GTACCATAACTAATAAAGCAGGGTACAGATATAAACTACTGCATCTTTTCTATAAACTGTGATTAAGAATTCTA
 CCTCTCCTGTATGGCTGTTACTGTACTGTACTCTGACTCCTTACCTAACAAATGAATTTGTTACATAATCTTCT
 ACATGTATGATTTGTGCCACTGATCTTAAACCTATGATTCAGTAACCTCTTACCATATAAAAACGATAATTGCTT
 TATTTGGAAGAATTTAGGAATACTAAGGACAATTTATTTTATAGACAAAGTAAAAAGACAGATATTTAAGAGG
 CATAACCAAAAAAGCAAACTTGTAAACAGAGTAAAAATCTTTAATATTTCTAAAGACATACTGTTTATCTGCTT
 CATATGCTTTTTTTAATTTCACTATTCCATTTCTAAATTAAGTTATGCTAAATGAGTAAGCTGTTTATCACTT
 AACAGCTCATTTTGTCTTTTCAATATACAAATTTTAAAAATACTACAATATTTAACTAAGGCCCAACCGATTTC
 CATAATGTAGCAGTTACCGTGTTCACCTCACACTAAGGCCTAGAGTTTGCTCTGATATGCATTTGGATGATTAAT
 GTTATGCTGTTCTTTTCATGTGAATGTCAAGACATGGAGGGTGTGTAATTTTATGGTAAATTAATCCTTCTTA
 CACATAATGGTGTCTTAAATTTGACAAAAAATGAGCACTTACAATTTGATGTCTCCTCAAATGAAGATTCTTTAT
 GTGAAATTTTAAAGACATTGATTCCGCATGTAAGGATTTTTCATCTGAAGTACAATAATGCACAATCAGTGTTG
 CTCAAACCTGCTTTTATACTTATAAACAGCCATCTTAAATAAGCAACGTATTGTGAGTACTGATATGTATATAATA
 AAATTATCAAAGGAAAA

FIGURE 181

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA52196

><subunit 1 of 1, 229 aa, 1 stop

><MW: 26017, pI: 4.73, NX(S/T): 0

MGDKIWLPFPVLLLLAALPPVLLPGAAGFTPSLSDSDFTFLLPAGQKECFYQPMPLKASLEIEY
QVLDGAGLDIDFHLASPEGKTLVFEQRKSDGVHTVETEVGDYMFCDNTFSTISEKVIFEL
ILDNMGEQAQEQEDWKYITGTDILDMKLEDILESINSIKSRLSKSGHIQILLRAFEARDN
IQESNFDRVNFWSMVNLVVMVVVSAIQVYMLKSLFEDKRKSRT

Important features:

Signal peptide:

amino acids 1-23

Transmembrane domain:

amino acids 195-217

N-myristoylation site.

amino acids 43-48

Tyrosine kinase phosphorylation site.

amino acids 55-62

FIGURE 182

CCATCCCTGAGATCTTTTTATAAAAAACCCAGTCTTTGCTGACCAGACAAAGCATACCAGAT
CTCACCAGAGAGTCGCAGACACTATGCTGCCTCCCATGGCCCTGCCCAGTGTGTCCTGGATG
CTGCTTTCCTGCCTCATTCTCCTGTGTCAGGTTCAAGGTGAAGAAACCCAGAAGGAACTGCC
CTCTCCACGGATCAGCTGTCCCAAAGGCTCCAAGGCCTATGGCTCCCCCTGCTATGCCTTGT
TTTTGTACCCAAAATCCTGGATGGATGCAGATCTGGCTTGCCAGAAGCGGCCCTCTGGAAAA
CTGGTGTCTGTGCTCAGTGGGGCTGAGGGATCCTTCGTGTCCTCCCTGGTGAGGAGCATTAG
TAACAGCTACTCATACATCTGGATTGGGCTCCATGACCCACACAGGGCTCTGAGCCTGATG
GAGATGGATGGGAGTGGAGTAGCACTGATGTGATGAATTACTTTGCATGGGAGAAAAATCCC
TCCACCATCTTAAACCCCTGGCCACTGTGGGAGCCTGTCAAGAAGCACAGGATTTCTGAAGTG
GAAAGATTATAACTGTGATGCAAAGTTACCCCTATGTCTGCAAGTTCAAGGACTTAGGGGCAGGT
GGGAAGTCAGCAGCCTCAGCTTGGCGTGCAGCTCATCATGGACATGAGACCAGTGTGAAGAC
TCACCCTGGAAGAGAATATTCTCCCCAACTGCCCTACCTGACTACCTTGTGTCATGATCCTCC
TTCTTTTTCCTTTTTCTTCACCTTCATTTACAGGCTTTTCTCTGTCTTCCATGTCTTGAGATC
TCAGAGAATAATAATAAAAATGTTACTTTATAAAAAAAAAAAAAAAAAAAAAA

FIGURE 183

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56965

<subunit 1 of 1, 175 aa, 1 stop

<MW: 19330, pI: 7.25, NX(S/T): 1

MLPPMALPSVSWMLLSCLILLCQVQGEETQKELPSPRISCPKGSKAYGSPCYALFLSPKSWM
DADLACQKRPSGKLVSVLGAEGSFVSSLVRSISNSYSYIWIGLHDPTQGSEPDGDGWEWSS
TDVMNYFAWEKNPSTILNPGHCGSLSRSTGFLKWKDYNCDAKLPYVCKFKD

Important features:

Signal peptide:

amino acids 1-26

C-type lectin domain signature.

amino acids 146-171

FIGURE 183

FIGURE 184

CCAGTCTGTCGCCACCTCACTTGGTGTCTGCTGTCCCCGCCAGGCAAGCCTGGGGTGAGAGC
ACAGAGGAGTGGGCCGGGACCATGCGGGGGACGCGGCTGGCGCTCCTGGCGCTGGTGCTGGC
TGCCTGCGGAGAGCTGGCGCCGGCCCTGCGCTGCTACGTCTGTCCGGAGCCCACAGGAGTGT
CGGACTGTGTCACCATCGCCACCTGCACCACCAACGAAACCATGTGCAAGACCACACTCTAC
TCCCGGGAGATAGTGTACCCCTTCCAGGGGGACTCCACGGTGACCAAGTCCTGTGCCAGCAA
GTGTAAGCCCTCGGATGTGGATGGCATCGGCCAGACCCTGCCCGTGTCTTGCTGCAATACTG
AGCTGTGCAATGTAGACGGGGCGCCCGCTCTGAACAGCCTCCACTGCGGGGGCCCTCACGCTC
CTCCCACTCTTGAGCCTCCGACTGTAGAGTCCCCGCCACCCCCATGGCCCTATGCGGCCCA
GCCCCGAATGCCTTGAAGAAGTGCCCCCTGCACCAGGAAAAAAAAAAAAAAAAAAAA

FIGURE 184

FIGURE 185

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56405

<subunit 1 of 1, 125 aa, 1 stop

<MW: 13115, pI: 5.90, NX(S/T): 1

MRGTRLALLALVLAACGELAPALRCYVCPEPTGVSDCVTIATCTTNETMCKTTLYSREIVYP

FQGDSTVTKSCASKCKPSDVDGIGQTLPVSCCNTELCNVDGAPALNSLHCGALTLLPLLSRL

Important features:

Signal peptide:

amino acids 1-17

N-glycosylation site.

amino acids 46-49

TopHat 2.9.6.0001

FIGURE 186

CTGCAGTCAGGACTCTGGGACCGCAGGGGGCTCCCGGACCCTGACTCTGCAGCCGAACCGGC
ACGGTTTCGTGGGGACCCAGGCTTGCAAAGTGACGGTCATTTTCTCTTTCTTTCTCCCTCTT
GAGTCCTTCTGAGATGATGGCTCTGGGCGCAGCGGGAGCTACCCGGGTCTTTGTGCGGATGG
TAGCGGCGGCTCTCGGCGGCCACCCTCTGCTGGGAGTGAGCGCCACCTTGAACCTCGGTTCTC
AATTCCAACGCTATCAAGAACCTGCCCCACCGCTGGGCGGCGCTGCGGGGCACCCAGGCTC
TGCAGTCAGCGCCGCGCCGGGAATCCTGTACCCGGGCGGGAATAAGTACCAGACCATTGACA
ACTACCAGCCGTACCCGTGCGCAGAGGACGAGGAGTGCGGCACTGATGAGTACTGCGCTAGT
CCCACCCGCGGAGGGGACGCAGGCGTGCAAATCTGTCTCGCCTGCAGGAAGCGCCGAAAACG
CTGCATGCGTCACGCTATGTGCTGCCCCGGGAATTACTGCAAAAATGGAATATGTGTGTCTT
CTGATCAAAATCATTTCCGAGGAGAAATTGAGGAAACCATCACTGAAAGCTTTGGTAATGAT
CATAGCACCTTGGATGGGTATTCCAGAAGAACCACCTTGTCTTCAAAAATGTATCACACCAA
AGGACAAGAAGGTTCTGTTTGTCTCCGGTCATCAGACTGTGCCTCAGGATTGTGTTGTGCTA
GACACTTCTGGTCCAAGATCTGTAAACCTGTCTGAAAGAAGGTCAAGTGTGTACCAAGCAT
AGGAGAAAAGGCTCTCATGGACTAGAAATATTCCAGCGTTGTTACTGTGGAGAAGGTCTGTC
TTGCCGGATACAGAAAGATCACCATCAAGCCAGTAATTCTTCTAGGCTTCACACTTGTGAGA
GACACTTAAACCAGCTATCCAAATGCAGTGAACCTCTTTTATATAATAGATGCTATGAAAACC
TTTTATGACCTTCATCAACTCAATCCTAAGGATATACAAGTTCTGTGGTTTTAGTTAAGCAT
TCCAATAACACCTTCCAAAAACCTGGAGTGTAAGAGCTTTGTTTCTTTATGGAACCTCCCCTG
TGATTGCAGTAAATTAAGTATTGTAAATTTCTCAGTGTGGCACTTACCTGTAAATGCAATGA
AACTTTTAATTATTTTTCTAAAGGTGCTGCACTGCCTATTTTTCTCTTGTATGTAAATTT
TTGTACACATTGATTGTTATCTTGACTGACAAATATTCTATATTGAACTGAAGTAAATCATT
TCAGCTTATAGTTCTTAAAAGCATAACCCTTTACCCCATTTAATTCTAGAGTCTAGAACGCA
AGGATCTCTTGGAATGACAAATGATAGGTACCTAAAATGTAACATGAAAATACTAGCTTATT
TTCTGAAATGTACTATCTTAATGCTTAAATTATATTTCCCTTTAGGCTGTGATAGTTTTTGA
AATAAAATTTAACATTTAAAAA

FIGURE 187

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA57530

<subunit 1 of 1, 266 aa, 1 stop

<MW: 28672, pI: 8.85, NX(S/T): 1

MMALGAAGATRVFVAMVAAALGGHPLLGVSATLNSVLNSNAIKNLPPPLGGAAGHPGSAVSA
APGILYPGGNKYQTIDNYQPYPCAEDEECGTDEYCASPTRGGDAGVQICLACRKRRCMRH
AMCCPGNYCKNGICVSSDQNHFRGEIEETITESFGNDHSTLDGYSRRTTLSSKMYHTKGQEG
SVCLRSSDCASGLCCARHFWSKICKPVLKEGQVCTKHRRKGSHGLEIFQRCYCGEGLSCRIQ
KDDHQASNSSLHTCQRH

Important features:

Signal peptide:

amino acids 1-23

N-glycosylation site.

amino acids 256-259

Fungal Zn(2)-Cys(6) binuclear cluster domain

amino acids 110-126

[illegible][illegible]

Year	Age	Sex	Occupation	Education	Marital Status	Religion	Political Party	Income	Health	Smoking	Alcohol	Drugs	Other
1990	25	Male	Student	High School	Single	Catholic	Democrat	\$10,000	Good	Yes	No	No	No
1991	26	Female	Teacher	College	Married	Protestant	Republican	\$15,000	Good	No	No	No	No
1992	27	Male	Engineer	College	Single	Jewish	Democrat	\$20,000	Good	No	No	No	No
1993	28	Female	Nurse	College	Married	Muslim	Democrat	\$18,000	Good	No	No	No	No
1994	29	Male	Doctor	College	Single	Hindu	Democrat	\$25,000	Good	No	No	No	No
1995	30	Female	Lawyer	College	Married	Buddhist	Democrat	\$30,000	Good	No	No	No	No
1996	31	Male	Artist	College	Single	Sikh	Democrat	\$12,000	Good	No	No	No	No
1997	32	Female	Writer	College	Married	Christian	Democrat	\$16,000	Good	No	No	No	No
1998	33	Male	Scientist	College	Single	Muslim	Democrat	\$22,000	Good	No	No	No	No
1999	34	Female	Manager	College	Married	Hindu	Democrat	\$19,000	Good	No	No	No	No
2000	35	Male	Analyst	College	Single	Buddhist	Democrat	\$21,000	Good	No	No	No	No
2001	36	Female	Consultant	College	Married	Sikh	Democrat	\$23,000	Good	No	No	No	No
2002	37	Male	Developer	College	Single	Christian	Democrat	\$24,000	Good	No	No	No	No
2003	38	Female	Designer	College	Married	Muslim	Democrat	\$26,000	Good	No	No	No	No
2004	39	Male	Researcher	College	Single	Hindu	Democrat	\$27,000	Good	No	No	No	No
2005	40	Female	Analyst	College	Married	Buddhist	Democrat	\$28,000	Good	No	No	No	No
2006	41	Male	Manager	College	Single	Sikh	Democrat	\$29,000	Good	No	No	No	No
2007	42	Female	Consultant	College	Married	Christian	Democrat	\$30,000	Good	No	No	No	No
2008	43	Male	Developer	College	Single	Muslim	Democrat	\$31,000	Good	No	No	No	No
2009	44	Female	Designer	College	Married	Hindu	Democrat	\$32,000	Good	No	No	No	No
2010	45	Male	Researcher	College	Single	Buddhist	Democrat	\$33,000	Good	No	No	No	No
2011	46	Female	Analyst	College	Married	Sikh	Democrat	\$34,000	Good	No	No	No	No
2012	47	Male	Manager	College	Single	Christian	Democrat	\$35,000	Good	No	No	No	No
2013	48	Female	Consultant	College	Married	Muslim	Democrat	\$36,000	Good	No	No	No	No
2014	49	Male	Developer	College	Single	Hindu	Democrat	\$37,000	Good	No	No	No	No
2015	50	Female	Designer	College	Married	Buddhist	Democrat	\$38,000	Good	No	No	No	No
2016	51	Male	Researcher	College	Single	Sikh	Democrat	\$39,000	Good	No	No	No	No
2017	52	Female	Analyst	College	Married	Christian	Democrat	\$40,000	Good	No	No	No	No
2018	53	Male	Manager	College	Single	Muslim	Democrat	\$41,000	Good	No	No	No	No
2019	54	Female	Consultant	College	Married	Hindu	Democrat	\$42,000	Good	No	No	No	No
2020	55	Male	Developer	College	Single	Buddhist	Democrat	\$43,000	Good	No	No	No	No
2021	56	Female	Designer	College	Married	Sikh	Democrat	\$44,000	Good	No	No	No	No
2022	57	Male	Researcher	College	Single	Christian	Democrat	\$45,000	Good	No	No	No	No
2023	58	Female	Analyst	College	Married	Muslim	Democrat	\$46,000	Good	No	No	No	No
2024	59	Male	Manager	College	Single	Hindu	Democrat	\$47,000	Good	No	No	No	No
2025	60	Female	Consultant	College</									

GAGGAACCTACCGGTACCGGCCGCGCGCTGGTAGTCGCCGGTGTGGCTGCACCTACCAATCCCGTGCGCCGCGG
 CTGGGCGCGTCGGAGAGTGCCTGTGCTTCTCTCCTGCACGCGGTGCTTGGGCTCGGCCAGGCGGGGTCCGCCGCCA
 GGGTTTGGAGGATGGGGGAGTAGCTACAGGAAGCGACCCCGCGATGGCAAGGTATATTTTTGTGGAATGAAAAGGA
 AGTATTAGAAATGAGCTGAAGACCATTCACAGATTAATATTTTTGGGGACAGATTTGTGATGCTTGATTCACCTT
 TGAAGTAATGTAGACAGAAGTTCTCAAATTTGCATATACATCAACTGGAACACAGCAGTGAATCTTAATGTTTAC
 TTAATACAGAACTTGCATATAAGAAAGAGAAATGGGAGTCTGGTTAAATAAGAGTAGACTATATCAGAGACTTTGAAAAG
 GATCATTTCTCTGTTTTCTGATAGTGATATGGCCATTTTAGTGGGCACAGATCAGGATTTTTACAGTTTACTTTGG
 AGTGTCCAAAACTGCAAGCAGTAGAGAAATAAGACAAGCTTTCAAGAAATTTGGCATTTGAAGTTACATCCTGATAA
 AAACCCGAATAACCCAAATGCACATGGCGATTTTTTAAAAATAAATAGAGCATATGAAGTACTCAAAGATGAAGA
 TCTACGGAAAAAGTATGACAAATATGGAGAAAAGGGACTTGAGGATAATCAAGGTGGCCAGTATGAAAGCTGGAA
 CTATATCTCGTTATGATTTTGGTATTTATGATGATGATCCTGAAATCATAACATTTGAAAGAAGAGAATTTGATG
 TGCTGTTTAATTTCTGGAGAAGTGGTTTGTAAATTTTTTACTCCCCAGGCTGTTTACACTGCCATGATTTAGCTCC
 CACATGGAGAGACTTTTGCATAAGAAAGTGGATGGGTTACTTTCGAATTTGGAGCTGTTAACTGTGGTGATGATAGAAT
 GCTTTGCCGAATGAAAGGAGTCAACAGCTATCCCAGTCTCTTCATTTTTCGGTCTGGAATGGCCCCAGTGAAATA
 TCATGGAGACAGATCAAAGGAGAGTTTAGTGAGTTTTTGCAATGCAGCATGTTAGAAGTACAGTGACAGAACTTTG
 GACAGGAAATTTTTGTCAACTCCATACAAACTGCTTTTTGTCTGCTGGTATTGGCTGGCTGATCACTTTTTTGTTCAAA
 AGGAGGAGATTTGTTGACTTCACAGACACGACTCAGGCTTTAGTGGCATGTTGTTTCTCAACTCATTTGGATGCTAA
 AGAAATATATTTGGAAGTAATACATAATCTTCCAGACTTTTGAACACTTTTTCGGCAAAACACTAGAGGATCGTTT
 GGCTCATCATCGGTGGCTGTTATTTTTTCACTTTTGGAAAAAATGAAAATTCAAATGATCCTGAGCTGAAAAAACT
 AAAAACTCTACTTAAAAATGATCATATTTCAAGTTGGCAGGTTTGACTGTTTCTCTGCACCAGACATCTGTAGTAA
 TCTGTATGTTTTTTCAGCCGTCTCTAGCAGTATTTAAAGGACAAGGAACCAAAGAAATATGAAATTCATCATGGAAA
 GAAGATTTCTATATGATATACTTTGCCCTTTGCCAAAGAAAGTGTGAATTTCTATGTTTACCACGCTTGGACCTCAAAA
 TTTTCTGCCAATGACAAAGAACCATGGCTGTTGATTTCTTTGCCCCCTGGTGTCCACCATGTCGAGCTTTACT
 ACCAGAGTTACGAAGAGCATCAAATCTTCTTTATGTTGAGCTTAAAGTTTGGTACACTAGATTTGTACAGTTTATGA
 GGGACTCTGTAACATGTATAACATTCAGGCCTTATCCAACAACAGTGGTATTCAACCAGTCCAACATTCATGAGTA
 TGAAGGACATCACTCTGCTGAACAAATCTTGGAGTTCATAGAGGATCTTATGAATCCTTCAGTGGTCTCCCTTAC
 ACCCACCACCTTCAACGAACCTAGTTACACAAAGAAAACACAAACGAAGTCTGGATGGTTGATTTCTATTCTCCGT
 GTGTCACTCTTGCCAAAGTCTTAATGCCAGAAATGGAAAAAATGGCCCCGACATTAACCTGGACTGATCAACGTGGG
 CAGTATAGATTTGCCAACAGTATCATTTCTTTTTTGCCCCAGGAAAACGTTCAAAGATAACCTCGAGATAAGATTTTT
 TCCCTCAAAATCAAATAAAGCTTATCAGTATCACAGTTACAATGGTTGGAATAGGGATGCTTTATTTCCCTGAGAA
 CTGGGGTCTAGGATTTTTTACCTCAAGTATCCACAGATCTAACACCTCAGACTTTCAGTGAAAAAGTTCTACAAGG
 GAAAAATCATTTGGGTGATTGATTTCTATGCTCCTTGGTGTGGACCTTGCCAGAATTTTGTCTCCAGAATTTGAGCT
 CTTGGCTAGGATGATTTAAAGGAAAAGTGAAAGCTGGAAAAGTAGACTGTCAGGCTTATGCTCAGACATGCCAGAA
 AGCTGGGATCAGGGCTTATCCAACCTGTTAAGTTTTATTTCTACGAAAGAGCAAAGAGAAAAATTTCAAGAAGAGCA
 GATAAATACAGAGATGCAAAAGCAATCGCTGCCCTTAATAAGTGA AAAATTTGAAAACTTCCGAAATCAAGGCAA
 GAGGAATAAGGATGAACCTTTGATAATGTTGAAGATGAAGAAAAAGTTTAAAAAGAAATTTTGACAGATGACATCAG
 AAGACACCTATTTAGAATGTTACATTTATGATGGGAATGAATGAACATTTATCTTAGACTTGCAGTTGTACTGCCA
 GAATTTATCTACAGCACTGGTGTAAAAGAAGGGTCTGCAAACTTTTTCTGTAAAGGGCCGGTTTTATAAATATTTTA
 GACTTTGTCAGGCTATAATATATGGTTTACACATGAGAACAAAGAATAGAGTCATCATGTAATCTTTGTTATTTGCT
 TTTAAACAACCTTTAAAAAATATTTAAACGATTTCTTAGCTCAGAGCCATACAAAAGTAGGCTGGATTCAGTCCATG
 GACCATAGATTTGCTGTCCCCCTCGACGGACTTATAATGTTTCAGGTGGCTGGCTTGAACATGAGTCTGCTGTGCT
 ATCTACATAAATGCTCTAAGTTGTATAAAGTCCACTTTTCCCTTCACGTTTTTTGGCTGACCTGAAAAGAGGTAAC
 TAGTTTTTTGGTCACTTGTCTCTCTAAAAATGCTATCCCTAACCATATATTTATATTTTCGTTTTAAAAACACCCAT
 GATGTGGCACAGTAAACAAACCTGTTATGCTGTATTATTATGAGGAGATTCTTCATTGTTTTCTTTCTTCTCA
 AAGGTTGAAAAAATGCTTTTAATTTTTTACAGCCGAGAAAACAGTGCAGCAGTATATGTGCACACAGTAAGTACAC
 AAATTTGAGCAACAGTAAGTGCACAAATTTCTGTAGTTTTGCTGTATCATCCAGGAAAAACCTGAGGGAAAAAATTA
 TAGCAATTAACCTGGCATTTGATAGATATCTTAATATGTTATCAAGTATTAGAGTTCTATATTTTAAAGATATA
 TGTGTTTCATGTATTTTCTGAAATGCTTTTATAGAAATTTTCCCACTGATAGTTGATTTTTTGAGGCATCTAATAT
 TTACATATTTGCCTTCTGAACTTTGTTTTGACCTGTATCCTTTATTTACATTTGGGTTTTTTCTTTTATAGTTTGTG
 TTTTTCACTCCTGTCCAGTCTATTTATTTATTTCAAATAGGAAAAATTACTTTACAGGTTGTTTTACTGTAGCTTAT
 AATGATACTGTAGTTATTTCCAGTTACTAGTTTACTGTCTAGAGGGCTGCCTTTTTTCAGATAAATATTTGACATAATA
 ACTGAAAGTATTTTTATAAGAAAAATCAAGTATATAAAATCTAGGAAAGGGATCTTCTAGTTTTCTGTGTTGTTTGA
 CTCAAAGAAATCAAAATTTGTAGTAACTGTAGTTGTTTGTATATAATTTAGTTTACAGATGTACAGAATGTTAAATTT
 CCAATCAGTCAAAAGAGGTCAATGAATTTAAAGGCTTGCACCTTTTTCAAAAAAATTTTTAAAAA

FIGURE 190

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56439

<subunit 1 of 1, 747 aa, 1 stop

<MW: 86127, pI: 7.46, NX(S/T): 2

MGVWLNKDDYIRDLCRIILCFLIVYMAILVGTDQDFYSLLGVSKTASSREIRQAFKKLALKL
HPDKNPNNPNAHGDFLKINRAYEVLKDEDLRKKYDKYGEKGLLEDNQGGQYESWNYRYDFGI
YDDDPEIITLERREFDAAVNSGELWVFVNFYSPGCSHCHDLAPTWRDFAKEVDGLLRIGAVNC
GDDRMLCRMKGVNSYPSLFI FRSGMAPVKYHGDRSKESLVSFAMQHVRSTVTELWTGNFVNS
IQTAFAAGIGWLITFC SKGGDCLTSQTRLRLSGMLFLNSLDAKEIYLEVIHNLPDFELLSAN
TLEDRLAHHRWLLFFHFGKNENSNDPELKKLKTLLKNDHIQVGRFDCSSAPDICSNLYVFQP
SLAVFKGQGTKEYEIIHHGKKILYDILAFAKESVNSHVTTLGPNFPANDKEPWLVDFFAPWC
PPCRALLPELRRASNLLYGQLKFGTLDCTVHEGLCNMYNIQAYPTTVVFNQSNIEYEGHHS
AEQILEFIEDLMNP SVVSLTPTTFNELVTQRKHNEVWMVDFYSPWCHPCQVLMPEWKRMART
LTGLINVGSIDCQQYHSFCAQENVQRYPEIRFFPPKSNKAYQYHSYNGWNRDAYSLRIWGLG
FLPQVSTDLTPTQTFSEKVLQGNHWVIDFYAPWCGPCQNFAPPEFELLARMIKGVKAGKVDC
QAYAQTCQKAGIRAYPTVKFYFYERAKRNFQEEQINTRDAKAI AALISEKLETLRNQGRNKDEL

Important features:

Endoplasmic reticulum targeting sequence.

amino acids 744-747

Cytochrome c family heme-binding site signature.

amino acids 158-163

Nt-dnaJ domain signature.

amino acids 77-96

N-glycosylation site.

amino acids 484-487

[illegible]

AGACAGTACCTCCTCCCTAGGACTACACAAGGACTGAACCCAGAAGGAAGGACAGAGCAAA
GCCATGAAACATCATCCTAGAAATCCTTCTGCTTCTGATCACCATCATCTACTCCTACTTGGAG
GTCGTTGGTGAAGTTTTTTCATTCCCTCAGAGGAGAAAATCTGTGGCTGGGGAGATTGTTCTCA
TTACTGGAGCTGGGCATGGAATAGGCAGGCAGACTACTTATGAATTTGCAAAACGCACAGAGC
ATATTGGTTCTGTGGGATATTAATAAGCGCGGTGTGGAGGAACTGCAGCTGAGTGCCGAAA
ACTAGGCGTCACTGCGCATGCGTATGTGGTAGACTGCAGCAACAGAGAAGAGATCTATCGCT
CTCTAAATCAGGTGAAGAAAGAAGTGGGTGATGTAACAATCGTGGTGAATAATGCTGGGACA
GTATATCCAGCCGATCTTCTCAGCACCAAGGATGAAGAGATTACCAAGACATTTGAGGTCAA
CATCCTAGGACATTTTTTGGATCACAAAAGCACTTCTTCCATCGATGATGGAGAGAAATCATG
GCCACATCGTCACAGTGGCTTCAGTGTGCGGCCACGAAGGGATTCCCTTACCTCATCCCATAT
TGTTCCAGCAAATTTGCCGCTGTTGGCTTTCACAGAGGTCTGACATCAGAACTTCAGGCCTT
GGGAAAAAAGTGGTATCAAAACCTCATGTCTCTGCCCAGTTTTTGTGAATACTGGGTTCCACCA
AAAATCCAAGCACAAAGATTATGGCCTGTATTGGAGACAGATGAAGTCGTAAGAAGTCTGATA
GATGGAATACTTACCAATAAGAAAATGATTTTTTGTTCATCGTATATCAATATCTTTCTGAG
ACTACAGAAGTTTCTTCCTGAACGCGCCTCAGCGATTTTAAATCGTATGCAGAATATTCAAT
TTGAAGCAGTGGTTGGCCACAAAATCAAAATGAAATTGAATAAATAAGCTCCAGCCAGAGATG
TATGCATGATAATGATATGAATAGTTTTCGAATCAATGCTGCAAAGCTTTATTTTACATTTTTT
TCAGTCCTGATAATATTA AAAACATTGGTTTTGGCACTAGCAGCAGTCAAACGAACAAGATTA
ATTACCTGTCTTCCTGTTTTCTCAAGAATATTTACGTAGTTTTTTCATAGGTCTGTTTTTCTT
TCATGCCTCTTAAAAACTTCTGTGCTTACATAAACATACTTAAAAGGTTTTCTTTAAGATAT
TTTATTTTTTCCATTTAAAGGTGGACAAAAGCTACCTCCCTAAAAGTAAATACAAAGAGAACT
TATTTACACAGGGAAGGTTTAAGACTGTTCAAGTAGCATTCCAATCTGTAGCCATGCCACAG
AATATCAACAAGAACACAGAATGAGTGCACAGCTAAGAGATCAAGTTTCAGCAGGCAGCTTT
ATCTCAACCTGGACATATTTTAAGATTTCAGCATTTGAAAGATTTCCCTAGCCTCTTCCTTTT
TCATTAGCCCAAACGGTGCAACTCTATTCTGGACTTTATTACTTGATTCTGTCTTCTGTAT
AACTCTGAAGTCCACCAAAGTGGACCCTCTATATTTCTCCCTTTTTTATAGTCTTATAAGA
TACATTATGAAAGGTGACCGACTCTATTTTAAATCTCAGAATTTTAAGTTCTAGCCCCATGA
TAACCTTTTTTCTTTGTAATTTATGCTTTCATATATCCTTGGTCCCAGAGATGTTTAGACAAT
TTTAGGCTCAAAAATTAAAGCTAACACAGGAAAAGGAACTGTACTGGCTATTACATAAGAAA
CAATGGACCCAAGAGAAGAA

FIGURE 192

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56409

<subunit 1 of 1, 300 aa, 1 stop

<MW: 33655, pI: 9.31, NX(S/T): 1

MNIILEILLLLITIIYSYLESLVKFFIPQRRKSVAGEIVLITGAGHGIGRQTTYEFAKRQSI
LVLWDINKRGVEETAAECKRLGVTAHAYVVDCSNREEIYRSLNQVKKEVGDTVIVVNNAGTV
YPADLLSTKDEEITKTTFEVNILGHFWITKALLPSMMERNHGHIVTVASVCGHEGIPYLIPYC
SSKFAAVGFHRGLTSELQALGKTGIKTSCLCPVFNVTGFTKNPSTRLWPVLETDEVVRSLLID
GILTNNKMIFVPSYINIFLRLQKFLPERASAILNRMQNIQFEAVVGHKIKMK

Important features:

Signal peptide:

amino acids 1-19

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 30-33 and 58-61

Short-chain alcohol dehydrogenase family protein

amino acids 165-202, 37-49, 112-122 and 210-219

FIGURE 193

CGGCGGCGGCTGCGGGCGCGAGGTGAGGGGCGCGAGGTGAGGGGCGCGAGGTTCACAGCAGG
ATGCCCCGGCTCTGCAGGAAGCTGAAGTGAGAGGCCCGAGAGGGGCCAGCCCGCCGGGGC
AGGATGACCAAGGCCCGGCTGTTCCGGCTGTGGCTGGTGCTGGGGTCGGTGTTTCATGATCCT
GCTGATCATCGTGTAAGTGGGACAGCGCAGGCGCCGCGCACTTCTACTTGACACAGTCCTTCT
CTAGGCCGCACACGGGGCCCGCTGCCCACGCCCCGGGCCGACAGGGACAGGGAGCTCACG
GCCGACTCCGATGTCGACGAGTTTCTGGACAAGTTTCTCAGTGCTGGCGTGAAGCAGAGCGA
CCTTCCCAGAAAGGAGACGGAGCAGCCGCCTGCGCCGGGGAGCATGGAGGAGAGCGTGAGAG
GCTACGACTGGTCCCCGCGCGACGCCCCGCGCAGCCAGACCAGGGCCGGCAGCAGGCGGAG
CGGAGGAGCGTGCTGCGGGGCTTCTGCGCCAACTCCAGCCTGGCCTTCCCCACCAAGGAGCG
CGCATTTCGACGACATCCCCAACTCGGAGCTGAGCCACCTGATCGTGACGACCGGCACGGGG
CCATCTACTGCTACGTGCCAAGGTGGCCTGCACCAACTGGAAGCGCGTGATGATCGTGCTG
AGCGGAAGCCTGCTGCACCGCGGTGCGCCCTACCGCGACCCGCTGCGCATCCCGCGCGAGCA
CGTGACAACGCCAGCGCGCACCTGACCTTCAACAAGTTCTGGCGCCGCTACGGGAAGCTCT
CCCGCCACCTCATGAAGGTCAAGCTCAAGAAGTACACCAAGTTCCTCTTCGTGCGCGACCCC
TTCGTGCGCCTGATCTCCGCCTTCCGCAGCAAGTTCGAGCTGGAGAACGAGGAGTTCTACCG
CAAGTTCGCCGTGCCCATGCTGCGGCTGTACGCCAACACACCAGCCTGCCCGCCTCGGCGC
GCGAGGCCTTCCGCGCTGGCCTCAAGGTGTCCTTCGCCAACTTCATCCAGTACCTGCTGGAC
CCGCACACGGAGAAGCTGGCGCCCTTCAACGAGCACTGGCGGCAGGTGTACCGCCTCTGCCA
CCCGTGCCAGATCGACTACGACTTCGTGGGGAAGCTGGAGACTCTGGACGAGGACGCCGCGC
AGCTGCTGCAGCTACTCCAGGTGGACCGGCAGCTCCGCTTCCCCCGAGCTACCGGAACAGG
ACCGCCAGCAGCTGGGAGGAGGACTGGTTCGCCAAGATCCCCCTGGCCTGGAGGCAGCAGCT
GTATAAACTCTACGAGGCCGACTTTGTTCTCTTCGGCTACCCCAAGCCCCGAAAACCTCCTCC
GAGACTTGAAAGCTTTCGCGTTGCTTTTTCTCGCGTGCTGGAACCTGACGCACGCGCACTCC
AGTTTTTTTATGACCTACGATTTTGCAATCTGGGCTTCTTGTTCACTCCACTGCCTCTATCC
ATTGAGTACTGTATCGATATTGTTTTTTAAGATTAATATATTCAGGTATTTAATACGA

FIGURE 194

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56112

<subunit 1 of 1, 414 aa, 1 stop

<MW: 48414, pI: 9.54, NX(S/T): 4

MTKARLFRLWLVLGSVFMILLIIVYWDSAGAAHFYLHTSF SRPHTGPPLPTPGPDRDRELT
DSDVDEFLLDKFLSAGVKQSDLPRKETEQPPAPGSMEESVRGYDWSPRDARRSPDQGRQQAER
RSVLRGFCANSSLAFTKERAFDDIPNSELSHLI VDDRHGAIYCYVPKVACTNWK RVMIVLS
GSL LHRGAPYRDPLRIPREHVHNASAH LTFNKFWRRYGKLSRHLMKVKLK KYTKFLFVRDPF
VRLISAFRSKFELNEEFYRKFAVPMLRLYANHTSLPASAREAFRAGLKVSFANFIQYLLDP
HTEKLAPFNEHWRQVYRLCHPCQIDYDFVGKLETLD EDAQAQLLQLLQVDRQLRFPPSYRNRT
ASSWEEDWFAKIPLAWRQQLYKLYEADFVLFGYPK PENLLRD

Important features:

Signal peptide:

amino acids 1-31

N-glycosylation sites.

amino acids 134-137, 209-212, 280-283 and 370-373

TNFR/NGFR family cysteine-rich region protein

amino acids 329-332

FIGURE 195

TCGGGCCAGAATTCGGCACGAGGCGGCACGAGGGCGACGGCCTCACGGGGCTTTGGAGGTGA
AAGAGGCCCCAGAGTAGAGAGAGAGAGAGACCGACGTACACGGGATGGCTACGGGAACGCGCT
ATGCCGGGAAGGTGGTGGTCGTGACCGGGGGCGGGCGCGGCATCGGAGCTGGGATCGTGCGC
GCCTTCGTGAACAGCGGGGCCCCGAGTGGTTATCTGCGACAAGGATGAGTCTGGGGGCCGGGC
CCTGGAGCAGGAGCTCCCTGGAGCTGTCTTTATCCTCTGTGATGTGACTCAGGAAGATGATG
TGAAGACCCTGGTTTCTGAGACCATCCGCCGATTTGGCCGCCTGGATTGTGTTGTCAACAAC
GCTGGCCACCACCCACCCCCACAGAGGCCTGAGGAGACCTCTGCCCAGGGATTCCGCCAGCT
GCTGGAGCTGAACCTACTGGGGACGTACACCTTGACCAAGCTCGCCCTCCCCTACCTGCGGA
AGAGTCAAGGGAATGTCATCAACATCTCCAGCCTGGTGGGGGCAATCGGCCAGGCCCAGGCA
GTTCCCTATGTGGCCACCAAGGGGGCAGTAACAGCCATGACCAAAGCTTTGGCCCTGGATGA
AAGTCCATATGGTGTCCGAGTCAACTGTATCTCCCCAGGAAACATCTGGACCCCGCTGTGGG
AGGAGCTGGCAGCCTTAATGCCAGACCCTAGGGCCACAATCCGAGAGGGCATGCTGGCCCAG
CCACTGGGCCCGCATGGGCCAGCCCGCTGAGGTGCGGGGCTGCGGCAGTGTTCTGGCCTCCGA
AGCCAACCTTCTGCACGGGCATTGAACTGCTCGTGACGGGGGGTGCGAGCTGGGGTACGGGT
GCAAGGCCAGTCGGAGCACCCCCGTGGACGCCCCCGATATCCCTTCCTTGATTTCTCTCATTT
CTACTTGGGGCCCCCTTCCTAGGACTCTCCCACCCCAAACCTCCAACCTGTATCAGATGCAGC
CCCCAAGCCCTTAGACTCTAAGCCCAGTTAGCAAGGTGCCGGGTCAACCTGCAGGTTCCCAT
AAAAACGATTTGCAGCC

FIGURE 196

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56045

<subunit 1 of 1, 270 aa, 1 stop

<MW: 28317, pI: 6.00, NX(S/T): 1

MATGTRYAGKVVVVTGGGRGIGAGIVRAFVNNSGARVVICDKDESGGRALEQELPGAVFILCD
VTQEDDVKTLVSETIRRFGRLCDVNNAGHHPPPQRPEETSAQGFRQLLELNLLGTYTLTKL
ALPYLRKSQGNVINISLVGAIGQAQAVPYVATKGAVTAMTKALALDESPYGVRVNCISPGN
IWTPLWEELAALMPDPRATIREGMLAQPLGRMGQPAEVGAAAVFLASEANFCTGIELLVTTGG
AELGYGCKASRSTPVDAPDIPS

Important features:

N-glycosylation site.

amino acids 138-141

Short-chain alcohol dehydrogenase family protein

amino acids 10-22, 81-91, 134-171 and 176-185

FIGURE 197

AGGCGGGCAGCAGCTGCAGGCTGACCTTGCAGCTTGGCGGAATGGACTGGCCTCACAACCTG
CTGTTTCTTCTTACCATTTCCATCTTCCTGGGGCTGGGCCAGCCCAGGAGCCCCAAAAGCAA
GAGGAAGGGGCAAGGGCGGCCTGGGCCCCTGGCCCCTGGCCCTCACCAGGTGCCACTGGACC
TGGTGTACGGATGAAACCGTATGCCCCGCATGGAGGAGTATGAGAGGAACATCGAGGAGATG
GTGGCCCAGCTGAGGAACAGCTCAGAGCTGGCCCAGAGAAAGTGTGAGGTCAACTTGCAGCT
GTGGATGTCCAACAAGAGGAGCCTGTCTCCCTGGGGCTACAGCATCAACCACGACCCCAGCC
GTATCCCCGTGGACCTGCCGGAGGCACGGTGCCCTGTGTCTGGGCTGTGTGAACCCCTTCACC
ATGCAGGAGGACCGCAGCATGGTGAGCGTGCCGGTGTTTCAGCCAGGTTCTGTGCGCCGCCG
CCTCTGCCCCGCCACCGCCCCGCACAGGGCCTTGCCGCCAGCGCGCAGTCATGGAGACCATCG
CTGTGGGCTGCACCTGCATCTTCTGAATCACCTGGCCCAGAAGCCAGGCCAGCAGCCCCGAGA
CCATCCTCCTTGCACCTTTGTGCCAAGAAAGGCCTATGAAAAGTAAACACTGACTTTTGAAA
GCAAG

FIGURE 198

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA59294

<subunit 1 of 1, 180 aa, 1 stop

<MW: 20437, pI: 9.58, NX(S/T): 1

MDWPHNLLFLLTISIFLGLGQPRSPKSKRKGQGRPGPLAPGPHQVPLDLVSRMKPYARMEY
ERNIEEMVAQLRNSSELAQRKCEVNLQLWMSNKRSLSPWGYSINHDPSRIPVDLPEARCLCL
GCVNPFTMQEDRSMVSVFVFSQVPVRRRLCPPPPRTGPCRQRAVMETIAVGCTCIF

Important features:

Signal peptide:

amino acids 1-20

N-glycosylation site.

amino acids 75-78

Homologous region to IL-17

amino acids 96-180.

FIGURE 199

GCGCCGCCAGGCGTAGGCGGGGTGGCCCTTGCGTCTCCCGCTTCCTTGAAAAACCGGCGGG
CGAGCGAGGCTGCGGGCCGGCCGCTGCCCTTCCCCCACTCCCCGCCGAGAAGCCTCGCTCG
GCGCCCAACATGGCGGGTGGGCGCTGCGGCCCCGAGCTAACGGCGCTCCTGGCCGCCTGGAT
CGCGGCTGTGGCGGCGACGGCAGGCCCCGAGGAGGCCGCGCTGCCGCCGGAGCAGAGCCGGG
TCCAGCCCATGACCGCCTCCAAC TGGACGCTGGTGATGGAGGGCGAGTGGATGCTGAAATTT
TACGCCCCATGGTGTCCATCCTGCCAGCAGACTGATT CAGAATGGGAGGCTTTTGCAAAGAA
TGGTGAAATACTTCAGATCAGTGTGGGGAAGGTAGATGTCATTCAAGAACCAGGTTTGAGTG
GCCGCTTCTTTGTCACCACTCTCC CAGCATTTTTTTCATGCAAAGGATGGGATATTCCGCCGT
TATCGTGGCCCAGGAATCTTCGAAGACCTGCAGAATTATATCTTAGAGAAGAAATGGCAATC
AGTCGAGCCTCTGACTGGCTGGAATCCCCAGCTTCTCTAACGATGTCTGGAATGGCTGGTC
TTTTTAGCATCTCTGGCAAGATATGGCATCTTCACAACTATTT CACAGTGACTCTTGGAATT
CCTGCTTGGTGTCTTATGTGTTTTTCGTCATAGCCACCTTGGTTTTTGGCCTTTTTATGGG
TCTGGTCTTGGTGGTAATATCAGAATGTTTCTATGTGCCACTTCCAAGGCATTTATCTGAGC
GTTCTGAGCAGAATCGGAGATCAGAGGAGGCTCATAGAGCTGAACAGTTGCAGGATGCCGAG
GAGGAAAAAGATGATTCAAATGAAGAAGAAAAACAAAGACAGCCTTGTAGATGATGAAGAAGA
GAAAGAAGATCTTGCGCATGAGGATGAAGCAGAGGAAGAAGAGGAGGAGGACAACCTTGGCTG
CTGGTGTGGATGAGGAGAGAAGTGAGGCCAATGATCAGGGGCCCCCAGGAGAGGACGGTGTG
ACCCGGGAGGAAGTAGAGCCTGAGGAGGCTGAAGAAGGCATCTCTGAGCAACCCTGCCCAGC
TGACACAGAGGTGGTGGAAAGACTCCTTGAGGCAGCGTAAAAGTCAGCATGCTGACAAGGGAC
TG TAGATTTAATGATGCGTTTTCAAGAATACACACCAAAACAATATGTCAGCTTCCCTTTGG
CCTGCAGTTTGTACCAAATCCTTAATTTTTTCTGAATGAGCAAGCTTCTCTTAAAGATGCT
CTCTAGTCATTTGGTCTCATGGCAGTAAGCCTCATGTATACTAAGGAGAGTCTTCCAGGTGT
GACAATCAGGATATAGAAAAACAAACGTAGTGTTGGGATCTGTTTGGGAGACTGGGATGGGAA
CAAGTTTCATTTACTTAGGGGT CAGAGAGTCTCGACCAGAGGAGGCCATTCC CAGTCCTAATC
AGCACCTTCCAGAGACAAGGCTGCAGGCCCTGTGAAATGAAAGCCAAGCAGGAGCCTTGGCT
CCTGAGCATCCCCAAAGTGTAACGTAGAAGCCTTG CATCCTTTTCTTGTTAAAGTATTTAT
TTTTGTCAAATTGCAGGAAACATCAGGCACCACAGTG CATGAAAAATCTTTCACAGCTAGAA
ATTGAAAGGGCCTTGGGTATAGAGAGCAGCTCAGAAGTCATCC CAGCCCTCTGAATCTCCTG
TGCTATGTTTTATTTCTTACCTTTAATTTTTTCCAGCATTTCCACCATGGGCATT CAGGCTCT
CCACACTCTTCACTATTATCTCTTGGTCAGAGGACTCCAATAACAGCCAGGTTTACATGAAC
TGTGTTTTGTTCACTCTGACCTAAGGGGTTTAGATAATCAGTAACCATAACCCCTGAAGCTGT
GACTGCCAAACATCTCAAATGAAATGTTGTGGCCATCAGAGACTCAAAGGAAGTAAGGATT
TTACAAGACAGATTAAAAAAAATTGTTTTGTCCAAAATATAGTTGTTGTTGATTTTTTTTT
AAGTTTTCTAAGCAATATTTTTTCAAGCCAGAAGTCCTCTAAGTCTTGCCAGTACAAGGTAGT
CTTGTTGAAGAAAAGTTGAATACTGTTTTGTTTTTCATCTCAAGGGGTTCCCTGGGTCTTGAAC
TACTTTAATAATAACTAAAAAACCACTTCTGATTTTCCCTCAGTGATGTGCTTTTGGTGAAA
GAATTAATGAACTCCAGTACCTGAAAGTGAAAGATTTGATTTTGTTC CATCTTCTGTAATC
TTCCAAAGAATTATATCTTTGTAAATCTCTCAATACTCAATCTACTGTAAGTACCCAGGGAG
GCTAATTTCTTT

FIGURE 200

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56433

<subunit 1 of 1, 349 aa, 1 stop

<MW: 38952, pI: 4.34, NX(S/T): 1

MAGGRCGPQLTALLAAWIAAVAATAGPEEAALPPEQSRVQPMTASNWTLVMEGEWMLKFYAP
WCPSCQQTDSEWEAFKNGEILQISVGKVDVIQEPGLSGRFFVTTLPAFFHAKDGI FRRYRG
PGIFEDLQNYILEKKWQSVEPLTGWKSPASLTMSGMAGLFSISGKIWHLHNYFTVTLGIPAW
CSYVFFVIATLVFGLFMGLVLVVISSECFYVPLPRHLSERSEQNRRSEEAHRAEQLQDAEEEEK
DDSNEEENKDSLVDDEEEKEDLGDEDEAEEDNLAAGVDEERSEANDQGPPGEDGV TRE
EVEPEEAEEGISEQPCPADTEVVEDSLRQRKSQHADKGL

Important features:

Signal peptide:

amino acids 1-22

Transmembrane domain:

amino acids 191-211

N-glycosylation site.

amino acids 46-49

Thioredoxin family proteins. (homologous region to disulfide isomerase)

amino acids 56-72

Flavodoxin proteins

amino acids 173-187

[illegible]

ATCTGGTTGAACACTACTCTTAAGCTTAATTTGTTAAACTCCGGTAAGTACCTAGCCCACATGATT
TGA CTCAGAGATTCTCTTTTGTCCACAGACAGT CATCTCAGGGGCAGAAAGAAAGAGCTCC
CAAATGCTATATCTATTTCAGGGGCTCTCAAGAACAAATGGAATATCATCCTGATTTAGAAAAT
TTGGATGAAGATGGATATACTCAATTACACTTCGACTCTCAAAGCAATACCAGGATAGCTGT
TGTTTCAGAGAAAGGATCGTGTGCTGCATCTCCTCCTTGGCGCCTCATTGCTGTAATTTTGG
GAATCCTATGCTTGGTAATACTGGTGATAGCTGTGGTCCTGGGTACCATGGGGGTTCTTTCC
AGCCCTTGTCTCTAATTGGATTATATATGAGAAGAGCTGTTATCTATT CAGCATGTCACT
AAATTCCTGGGATGGAAGTAAAAGACAATGCTGGCAACTGGGCTCTAATCTCCTAAAGATAG
ACAGCTCAAATGAATTGGGATTTATAGTAAAACAAGTGTCTTCCCAACCTGATAATTCATTT
TGGATAGGCCTTTCTCGGCCCCAGACTGAGGTACCATGGCTCTGGGAGGATGGATCAACATT
CTCTTCTAACTTATTTTCAGATCAGAACCACAGCTACCCAAGAAAACCCATCTCCAAATTGTG
TATGGATTACAGTGTCACTCATTTTATGACCAACTGTGTAGTGTGCCCTCATATAGTATTTGT
GAGAAGAAGTTTTCAATGTAAAGAGGAAGGGTGGAGAAGGAGAGAGAAATATGTGAGGTAGTA
AGGAGGACAGAAAACAGAACAGAAAAGAGTAACAGCTGAGGTCAAGATAAATGCAGAAAATG
TTTAGAGAGCTTGGCCAACCTGTAATCTTAACCAAGAAATTGAAGGGAGAGGCTGTGATTTCT
GTATTTGTCTGACCTACAGGTAGGCTAGTATTATTTTTCTAGTTAGTAGATCCCTAGACATGG
AATCAGGGCAGCCAAGCTTGAGTTTTTATTTTTTATTTATTTATTTTTTTGAGATAGGGTCT
CACTTTGTTACCCAGGCTGGAGTGCAGTGGCACAATCTCGACTCACTGCAGCTATCTCTCG
CTCAGCCCCCTCAAGTAGCTGGGACTACAGGTGCATGCCACCATGCCAGGCTAATTTTTGGTG
TTTTTTGTAGAGACTGGGTTTTGCCATGTTGACCAAGCTGGTCTCTAACTCCTGGGCTTAAG
TGATCTGCCCCGCTTGGCCTCCCAAAGTGCTGGGATTACAGATGTGAGCCACCACACCTGGC
CCCAAGCTTGAATTTTCATTCTGCCATTGACTTGGCATTTACCTTGGGTAAAGCCATAAGCGA
ATCTTAATTTCTGGCTCTATCAGAGTTGTTTCATGCTCAACAATGCCATTGAAGTGCACGGT
GTGTTGCCACGATTTGACCTCAACTCTAGCAGTATATCAGTTATGAACTGAGGGTGAAAT
ATATTTCTGAATAGCTAAATGAAGAAATGGGAAAAAATCTTCACCACAGTCAGAGCAATTTT
ATTATTTTCATCAGTATGATCATAATTATGATTATCATCTTAGTAAAAAAGCAGGAACTCCTA
CTTTTCTTTATCAATTAATAGCTCAGAGAGTACATCTGCCATATCTCTAATAGAATCTTT
TTTTTTTTTTTTTTTTTTTTTGTAGACAGAGTTTCGCTCTTGTTGCCAGGCTGGAGTGCAACGG
CACGATCTCGGCTCACCGCAACCTCCGCCCCCTGGGTTCAGCAATTCTCCTGCCTCAGCCT
CCCAAGTAGCTGGGATTACAGTCAGGCACCACCACACCCGGCTAATTTTGTATTTTCTTAGT
AGAGACAGGGTTTTCTCCATGTCGGT CAGGGTAGTCCCGAACTCCTGACCTCAAGTGATCTGC
CTGCCCTCGGCCTCCCAAGTGCTGGGATTACAGGCGTGAGCCACTGCACCCAGCCTAGAATCT
TGTATAATATGTAATTGTAGGGAAACTGCTCTCATAGGAAAGTTTTCTGCTTTTTTAAATACA
AAAATACATAAAAAATACATAAAATCTGATGATGAATATAAAAAAGTAACCAACCTCATTGGA
ACAAGTATTAACATTTTGGAAATATGTTTTATTAGTTTTTGTGATGTACTGTTTTACAATTTTT
ACCATTTTTTTTCAGTAATTACTGTAAAAATGGTATTATTGGAATGAAACTATATTTCTCATG
TGCTGATTTGTCTTATTTTTTTTCATACTTTCCCACTGGTGCTATTTTTTATTTCCAATGGATA
TTTCTGTATTACTAGGGAGGCATTTACAGTCCTCTAATGTTGATTAAATATGTGAAAAGAAAT
TGTACCAATTTTACTAAATTATGCAGTTTAAAAATGGATGATTTTATGTTATGTGGATTTTCAT
TTCAATAAAAAAAACTCTTATCAA

FIGURE 202

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA53912

<subunit 1 of 1, 201 aa, 1 stop

<MW: 22563, pI: 4.87, NX(S/T): 1

MEYHPDLENLDEDGYTQLHFDSQSNTRIAVVSEKGS CAASPPWRLIAVILGILCLVILVIAV
VLGTMGVLSSPCPPNWIIYEKSCYLFMSLSNSWDGSKRQCWQLGSNLLKIDSSNELGFIVKQ
VSSQPDNSFWIGLSRPQTEVPWLWEDGSTFSSNLFQIRTTATQENPSPNCVWIHVSVIYDQL
CSVPSYSICEKKFSM

Important features:

Type II transmembrane domain:

amino acids 45-65

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 197-200

N-myristoylation sites.

amino acids 35-40 and 151-156

Homologous region to LDL receptor

amino acids 34-67 and 70-200.

[illegible][illegible]

FIGURE 204

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA50921

<subunit 1 of 1, 693 aa, 1 stop

<MW: 77738, pI: 8.87, NX(S/T): 7

MTPQSLLQTTLFLLSLLFLVQGAHGRGHREDFRFSQQRNQTHRSSLHYKPTPDLRISIENSE
EALT VHAPFPAAHPASRSFPDPRGLYHFCLYWNRHAGRLHLLYGKRDFLLSDKASSLLCFQH
QEESLAQGPPLLATSVTSWWSPQNI SLPSAASF TFSFHSPHTAAHNASVDMCELKRD LQLL
SQFLKHPQKASRRPSAAPASQQ LQSLESKLTSVRFMGDMVSFEEDRINATVWKLQPTAGLQD
LHIHSRQEEEQSEIMEYSVLLPRTL FQRTKGRSGEAEKRLLL VDFSSQALFQDKNSSQVLGE
KVLGIVVQNTKVANLTEPVVLTFQHQLQPKNVT LQCVFWVEDPTLSSPGHWSSAGCETVRRE
TQTSCFCNHLTYFAVLMVSSVEVD AVHKHYLSLLSYVGCVV SALACLVTIAAYLC SRVPLPC
RRKPRDYTIKVHMNLLLAVFLLDTSFLLSEPVALTGSEAGCRASAI FLHFSLLTCLSWMGLE
GYNLYRLVVEVFGTYVPGYLLKLSAMGWGFPI FLVTLVALVDVDNYGPI ILAVHRTPEGVIY
PSMCWIRDSLVS YITNLGLFSLVFLFN MAMLATMVVQILRLRPHTQKWSHVL TLLGLSLVLG
LPWALIFFSFASGTFQLVVL YLFSIITSFQGF LIFIWYWSMRLQARGGPSPLKSN SDSARLP
ISSGSTSSSRI

Important features:

Signal peptide:

amino acids 1-25

Putative transmembrane domains:

amino acids 382-398, 402-420, 445-468, 473-491, 519-537, 568-590
and 634-657

Microbodies C-terminal targeting signal.

amino acids 691-693

cAMP- and cGMP-dependent protein kinase phosphorylation sites.

amino acids 198-201 and 370-373

N-glycosylation sites.

amino acids 39-42, 148-151, 171-174, 234-237, 303-306, 324-327
and 341-344

G-protein coupled receptors family 2 proteins

amino acids 475-504

[illegible][illegible]

FIGURE 206

CGGACGCGTGGGCGGACGCGTGGGCGGACGCGTGGGCGGACGCGTGGGCTGGTTTCAGGTCCAGGTTTTGCTTTGA
TCCTTTTCAAAAAGTGGAGACACAGAAGAGGGCTCTAGGAAAAAGTTTTGGATGGGATTATGTGGAACTACCCCT
GCGATTCTCTGCTGCCAGAGCAGGCTCGGCGCTTCCACCCAGTGCAGCCTTCCCCTGGCGGTGGTGAAAGAGAC
TCGGGAGTCGCTGCTTCCAAAGTGCCCGCCGTGAGTGAGCTCTCACCCAGTGCAGCCAAATGAGCCTCTTCGGGC
TTCTCCTGCTGACATCTGCCCTGGCCGGCCAGAGACAGGGGACTCAGGCGGAATCCAACCTGAGTAGTAAATTC
AGTTTTTCAGCAACAAGGAACAGAACGGAGTACAAGATCCTCAGCATGAGAGAAATTATTAAGTGTCTACTAATG
GAAGTATTCACAGCCCAAGGTTTTCTCATACCTTATCCAAGAAATACGGTCTTGGTATGGAGATTAGTAGCAGTAG
AGGAAAATGTATGGATACAACCTTACGTTTGTATGAAAGATTTGGGCTTGAAGACCCAGAAGATGACATATGCAAGT
ATGATTTTGTAGAAGTTGAGGAACCCAGTGTGAACTATATTAGGGCGCTGGTGTGGTTCTGGTACTGTACCAG
GAAAACAGATTTCTAAAGGAAATCAAATTAGGATAAGATTTGTATCTGATGAATATTTTCCTTCTGAACCAGGGT
TCTGCATCCACTACAACATTGTATGCCACAATTACAGAAGCTGTGAGTCCCTCAGTGTACCCCCCTTCAGCTT
TGCCACTGGACCTGCTTAATAATGCTATAACTGCCCTTTAGTACCTTGGAAAGACCTTATTCGATATCTTGAACCAG
AGAGATGGCAGTTGGACTTAGAAGATCTATATAGGCCAACTTGGCAACTTCTTGGCAAGGCTTTTGTTTTTGGAA
GAAAATCCAGAGTGGTGGATCTGAACCTTCTAACAGAGGAGGTAAGATTATACAGCTGCACACCTCGTAACTTCT
CAGTGTCCATAAGGGAAGAATAAGAGAACCGATACCATTTTCTGGCCAGGTTGTCTCCTGGTTAAACGCTGTG
GTGGGAACTGTGCCTGTTGTCTCCACAATTGCAATGAATGTCAATGTGTCCCAAGCAAAGTTACTAAAAAATACC
ACGAGGTCTTTCAGTTGAGACCAAGACCGGTGTGAGGGGATTGCACAAATCACTCACCGACGTGGCCCTGGAGC
ACCATGAGGAGTGTGACTGTGTGTGCAGAGGGAGCACAGGAGGATAGCCGCATCACACCAGCAGCTCTTGCCCA
GAGCTGTGCAGTGCAGTGGCTGATTCTATTAGAGAACGTATGCGTTATCTCCATCCTTAATCTCAGTTGTTTTGCT
TCAAGGACCTTTCATCTTCAGGATTTACAGTGCATTCTGAAAGAGGAGACATCAAACAGAATTAGGAGTTGTGCA
ACAGCTCTTTTGGAGAGGAGGCTTAAAGGACAGGAGAAAAGGCTTCAATCGTGGAAAGAAAATTAAATGTTGTAT
TAAATAGATCACAGCTAGTTTCAGAGTTACCATGTACGTATTCCACTAGCTGGGTTCTGTATTTTCAGTTCTTTT
GATACGGCTTAGGGTAATGTGAGTACAGGAAAAAACTGTGCAAGTGAGCACCTGATTCGGTTCCTTGTCTTAAC
TCTAAAGCTCCATGTCTGCGGCTTAAATCGTATAAAATCTGGATTTTTTTTTTTTTTTTTTTGCTCATATTCACAT
ATGTAAACCAGAACATTCTATGTACTACAAACCTGGTTTTTAAAAAGGAACTATGTTGCTATGAATTAAACTTGT
GTCATGCTGATAGGACAGACTGGATTTTTCATATTTCTTATTAATTTCTGCCATTTAGAAGAAGAGAACTACA
TTCATGGTTTGGAAAGAGATAAACCTGAAAAGAAGAGTGGCCTTATCTTCACTTTATCGATAAGTCAGTTTATTTG
TTTCATTGTGTACATTTTTTATATTCTCCTTTTGACATTATAACTGTTGGCTTTTCTAATCTTGTAAATATATCT
ATTTTTACCAAAGGTATTTAATATTCTTTTTTATGACAACCTTAGATCAACTATTTTTCAGCTTGGTAAATTTTTCT
AAACACAATTGTTATAGCCAGAGGAACAAAGATGATATAAAATATTGTTGCTCTGACAAAAATACATGTATTTCA
TTCTCGTATGGTGTAGAGTTAGATTAATCTGCATTTTAAAAAACTGAATTGGAATAGAATTGGTAAGTTGCAAA
GACTTTTTGAAAATAATTAAATTATCATATCTTCCATTCTGTTATTTGGGAGATGAAAAATAAAAGCAACTTATGA
AAGTAGACATTGAGATCCAGCCATTACTAACCTATTCTTTTGGGGAAATCTGAGCCTAGCTCAGAAAAACAT
AAAGCACCTTGAAGAAAGACTTGGCAGCTTCTGATAAAGCGTGTGCTGTGCTGTGCTGAGTAGGAACACATCCTATTTA
TTGTGATGTTGTGGTTTTTATTATCTTAAACTCTGTTCCATACACTTGTATAAATACATGGATATTTTTTATGTACA
GAAGTATGTCTCTTAACCAGTTCACTTATTGTACTCTGGCAATTTAAAAAGAAAATCAGTAAAAATATTTTGCTTGT
AAAATGCTTAATATNGTGCCTAGGTTATGTGGTGACTATTTGAATCAAAAATGTATTGAATCATCAAATAAAGA
ATGTGGCTATTTTGGGGAGAAAATTAAAAAAGGTTTAGGGATAACAGGGTAATGCGGCC

FIGURE 207

MSLFGLLLLT SALAGQRQGTQAESNLSSKFQFSSNKEQNGVQDPQHERIITVSTNGSIHSPR
FPHTYPRNTVLVWRLVAVEENVWIIQLTFDERFGLEDPEDDICKYDFVEVEEPSDGTILGRWC
GSGTVPGKQISKGNQIRIRFVSDEYFPSEPGFCIHYNIVMPQFTEAVSPSVLPPSALPLDLL
NNAITAFSTLEDLIRYLEPERWQLDLEDLYRPTWQLLGKAFVFGRKSRVVDLNLLTEEVRLY
SCTPRNFVSVSIREELKRTDTIFWPGCLLVKRCGGNCACCLHNCNECQCVPSKVTKKYHEVLQ
LRPKTGVRGLHKSLTDVALEHHHEECDVCVRGSTGG

Signal sequence:

amino acids 1-14

FIGURE 208

CCCATCTCAAGCTGATCTTGGCACCTCTCATGCTCTGCTCTCTTCAACCAGACCTCTACATTCCATTTTGGGAAGA
 AGACTAAAAATGGTGTTCCTAATGTGGACACTGAAGAGACAAATTCCTATCCTTTTAAACATAATCCTAATTTCC
 AAACCTCCTTGGGGCTAGATGGTTTCTAAACTCTGCCCTGTGATGTCACTCTGGATGTTCCAAAGAACCATGTG
 ATCGTGGACTGCACAGACAAGCATTTGACAGAAATTCCTGGAGGTATTCCCACGAACACCACGAACCTCACCCCTC
 ACCATTAACCACATACCAGACATCTCCCCAGCGTCCTTTACAGACTGGACCATCTGGTAGAGATCGATTTTCAGA
 TGCAACTGTGTACCTATTCCACTGGGGTCAAAAAACAACATGTGCATCAAGAGGCTGCAGATTAAACCCAGAAGC
 TTAGTGGACTCACTTATTTAAATCCCTTTACCTGGATGGAACCCAGCTACTAGAGATACCGCAGGGCCTCCCCG
 CCTAGCTTACAGCTTCTCAGCCTTGAGGCCAACACATCTTTTCCATCAGAAAAGAGAATCTAACAGAACTGGCC
 AACATAGAAATACTCTACCTGGGCCAAAACCTGTTATTATCGAAATCCTTGTTATGTTTCATATCAATAGAGAAA
 GATGCCTTCTAAACTTGACAAAGTTAAAAGTGCTCTCCCTGAAAGATAACAATGTACAGCCGTCCTCTACTGTT
 TTGCCATCTACTTTAACAGAACTATATCTCTACAACAACATGATTGCAAAAATCCAAGAAGATGATTTTAATAAC
 CTCAACCAATTACAAATTTCTTGACCTAAGTGGAATTGCCCTCGTTGTTATAATGCCCCATTTCTTGTGCGCCG
 TGTAATAATAATTCTCCCTACAGATCCCTGTAAATGCTTTTGATGCGCTGACAGAATTAAGTTTTACGTCTA
 CACAGTAACCTCTTTCAGCATGTGCCCCAAGATGGTTTAAAGAACATCAACAACTCCAGGAACCTGGATCTGTCC
 CAAAACCTTCTTGGCCAAAGAAATTGGGGATGCTAAATTTCTGCATTTTCTCCCCAGCCTCATCCAATTGGATCTG
 TCTTTCAATTTTGAACCTTCAGGTCTATCGTGCATCTATGAATCTATCACAAGCATTTTCTTCACTGAAAAGCCTG
 AAAATTCTGCGGATCAGAGGATATGTCTTTAAAGAGTTGAAAAGCTTTAACCTCTCGCCATTACATAATCTTCAA
 AATCTTGAAGTTCTTGATCTTGGCACTAACTTTATAAAAAATGCTAACCTCAGCATGTTTAAACAATTTAAAAGA
 CTGAAAGTCATAGATCTTTCAGTGAATAAAATATCACCTTCAGGAGATTCAAGTGAAGTTGGCTTCTGCTCAAAAT
 GCCAGAACCTCTGTAGAAAGTTATGAACCCAGGTCTTGAACAATTACATTATTTAGATATGATAAGTATGCA
 AGGAGTTGCAGATTCAAAAACAAGAGGCTTCTTTTATGCTCTGTTAATGAAAGCTGCTACAAGTATGGGCAGACC
 TTGGATCTAAGTAAAAATAGTATATTTTGTCAAGTCTCTGATTTTTCAGCATCTTTCTTTCTCAAATGCCTG
 AATCTGTGAGAAATCTCATTAGCCAACTCTTAATGGCAGTGAATTTCAACCTTTAGCAGAGCTGAGATATTTG
 GACTTCTCCAACAACCGGCTTGATTTACTCCATTCAACAGCATTTGAAGAGCTTCACAACTGGAAGTTCTGGAT
 ATAAGCAGTAATAGCCATTATTTTCAATCAGAAGGAATTACTCATATGCTAAACTTTACCAAGAACCTAAAGGTT
 CTGCAGAACTGATGATGAACGACAATGACATCTTCTCCACCAGCAGGACCATGGAGAGTGAGTCTCTTAGA
 ACTCTGGAATTCAGAGGAAATCACTTAGATGTTTTATGGAGAGAAGGTGATAACAGATACTTACAATTATTCAAG
 AATCTGCTAAAATTAGAGGAATTAGACATCTCTAAAATTTCCCTAAGTTTCTTGCCCTCTGGAGTTTTTGTGTT
 ATGCCCTCAAATCTAAAGAATCTCTCTTTGGCCAAAATGGGCTCAAATCTTTCAGTTGGAAGAACTCCAGTGT
 CTAAAGAACCTGGAACTTTGGACCTCAGCCACAACCAACTGACCACTGTCCCTGAGAGATTATCCAACCTGTTCC
 AGAAGCCTCAAGAATCTGATTCTTAAGAATAATCAAATCAGGAGTCTGACGAAGTATTTTCTACAAGATGCCTTC
 CAGTTGCGATATCTGGATCTCAGCTCAAATAAAATCCAGATGATCCAAAAGACCAGCTTCCAGAAAATGTCCTC
 AACAACTGGAAGATGTTGCTTTTGCATCATAATCGGTTTCTGTGCACCTGTGATGCTGTGTGGTTTGTCTGGTGG
 GTTAACCATAACGGAGGTGACTATTCTTACCTGGCCACAGATGTGACTTGTGTGGGGCCAGGAGCACACAAGGGC
 CAAAGTGTGATCTCCCTGGATCTGTACACCTGTGAGTTAGATCTGACTAACCTGATTCTGTTCTCACTTTCCATA
 TCTGTATCTCTCTTTCTCATGGTGTATGATGACAGCAAGTCACCTCTATTTCTGGGATGTGTGGTATATTTACCAT
 TTCTGTAAGGCCAAGATAAAGGGGTATCAGCGTCTAATATCACCAGACTGTTGCTATGATGCTTTTATTGTGTAT
 GACACTAAAGACCCAGCTGTGACCGAGTGGGTTTTGGCTGAGCTGGTGGCCAACTGGAAGACCCAAAGAGAGAAA
 CATTTTAAATTTATGTCTCGAGGAAAGGGACTGGTTACCAGGGCAGCCAGTTCTGGAAAACCTTTCCAGAGCATA
 CAGCTTAGCAAAAAGACAGTGTGTGTGATGACAGACAAGTATGCAAGACTGAAAATTTTAAGATAGCATTTTAC
 TTGTCCCATCAGAGGCTCATGGATGAAAAGTTGATGTGATTATCTTGATATTTCTTGAGAAGCCCTTTTCAAGAG
 TCCAAGTTCTCCAGCTCCGAAAAGGCTCTGTGGGAGTTCTGTCTTGAGTGGCCAAACAAACCCGCAAGCTCAC
 CCATACTTCTGGCAGTGTCTAAAGAACGCCCTGGCCACAGACAATCATGTGGCCTATAGTCAGGTGTTCAAGGAA
 ACGGTCTAGCCCTTCTTTGCAAAACACAACCTGCCCTAGTTTACCAAGGAGAGGCCTGGC

FIGURE 209

MVFPMWTLKRQILILFNIILISKLLGARWFPKTLPCDVTLDVPKNHVIVDCTDKHLTEIPGG
IPTNTTNLTLTINHIPDISPASFHRLDHLVEIDFRCNCVPIPLGSKNNMCIKRLQIKPRSFS
GLTYLKSLYLDGNQLLEIPQGLPPSLQLLSLEANNIFSIRKENLTELANIEILYLQNCYYR
NPCYVSYSEIEKDAFLNLTCLKVLSLKDNNVTAVPTVLPSTLTLEYLYNNMIAKIQEDDFNNL
NQLQILDLSGNCPRCYNAPFPCAPCKNNSPLQIPVNAFDALTELKVLRLHSNSLQHVPPRWF
KNINKLQELDLSQNFLAKEIGDAKFLHFLPSLIQLDLSFNFELQVYRASMNLSQAFSSLKSL
KILRIRGYVFKELKSFNLSPLHNLQNLEVLDTGTNFIKIANLSMFKQFKRLKVIDLSVNKIS
PSGDSSEVGFCSNARTSVESYEPQVLEQLHYFRYDKYARSCRFKNKEASFMSVNESCYKYGQ
TLDLSKNSIFFVKSSDFQHLSFLKCLNLSGNLISQTLNGSEFQPLAELRYLDFSNNRDLHLH
STAFEELHKLEVLDISSNSHYFQSEGITHMLNFTKNLKVLQKLMMNDNDISSSTSRTMESES
LRTLEFRGNHLDVLWREGDNRYLQLFKNLLKLEELDISKNSLSFLPSGVFDGMPPNLKNLSL
AKNGLKSFSWKKLQCLKNLETLDLSHNQLTTVPERLSNCSRSLKNLILKNNQIRSLTKYFLQ
DAFQLRYLDLSSNKIQMIQKTSFPENVLNNLKMLLLHHNRFLCTCAVWFVWVWNHTEVTIP
YLATDVTCVGPGAHKGQSVISLDLYTCELDLTNLILFSLISISVSLFLMVMMTASHLYFWDVW
YIYHFCKAKIKGYQRLISPDCCYDAFIVYDTKDPVTEWVLAELVAKLEDPREKHFNLCLEE
RDWLPGQPVLENLSQSIQLSKKTVFVMTDKYAKTENFKIAFYLSHQRLMDEKVDVIIILIFLE
KPFQKSKFLQLRKRLCGSSVLEWPTNPQAHFYFWQCLKNALATDNHVAYSQVFKETV

Signal sequence:

amino acids 1-26

Transmembrane domain:

amino acids 840-860

FIGURE 210

GGGTACCATTCTGCGCTGCTGCAAGTTACGGAATGAAAAATTAGAACAAACAGAAACATGGAACATGTTCTTTCT
AGTCGTCAATGCTGACCTGCATTTTCTGCTAATATCTGGTTCTGTGAGTTATGCGCCGAAGAAAAATTTTCTA
GAAGCTATCCTTGTGATGAGAAAAAGCAAAATGACTCAGTTATTGTCAGAGTGCGCAATCGTCCGACTACAGGAAG
TTCCCCAAACGGTGGGCAAATATGTGACAGAACTAGACCTGTCTGATAATTTTCATCACACACATAACGAATGAAT
CATTTCAAGGGCTGCAAAATCTCACTAAAATAAATCTAAACCACAACCCCAATGTACAGCACCAGAACGGAAATC
CCGGTATACAATCAAATGGCTTGAATATCACAGACGGGGCATTCCTCAACCTAAAAAACCTAAGGGAGTTACTGC
TTGAAGACAACAGTTACCCCAAATACCCCTCTGGTTTGGCCAGAGTCTTTGACAGAACTTAGTCTAATTCAAAAACA
ATATATACAACATAACTAAAGAGGGCATTTCAAGACTTATAAACTTGAAAAATCTCTATTTGGCCTGGAACGTCT
ATTTTAACAAAGTTTGGCGAGAAAACCTAACATAGAAGTGGAGTATTTGAAACGCTGACAAATTTGGAGTTGCTAT
CACTATCTTTCAATTCTCTTTTACACGTGCCACCCAAACTGCCAAGCTCCCTACGCAAACTTTTTCTGAGCAACA
CCCAGATCAAATACATTAGTGAAGAAGATTTCAAGGGATTGATAAATTTAACATTACTAGATTTAAGCGGGAAC
GTCCGAGGTGCTTCAATGCCCCATTTCCATGCGTGCCTTGTGATGGTGGTCTTCAATTAATATAGATCGTTTTG
CTTTTCAAACTTGACCCAACTTCGATACCTAAACCTCTCTAGCACTTCCCTCAGGAAGATTAATGCTGCCTGGT
TTAAAAATATGCCTCATCTGAAGGTGCTGGATCTTGAATCAACTATTTAGTGGGAGAAATAGTCTCTGGGGCAT
TTTTAACGATGCTGCCCCGCTTAGAAATACCTGACTTGTCTTTTAACTATATAAAGGGGAGTTATCCACAGCATA
TTAATATTTCCAGAAACTTCTCTAAACTTTTGTCTCTACGGGCATTGCAATTTAAGAGGTTATGTGTTCAGGAAC
TCAGAGAAGATGATTTCCAGCCCCCTGATGCAGCTTCCAACTTATCGACTATCAACTTGGGTATTAATTTTATTA
AGCAAATCGATTTCAAACCTTTTCAAATTTCTCCAATCTGGAAATTAATTTACTTGTGAGAAAACAGAATATCAC
CGTTGGTAAAGATACCCGGCAGAGTTATGCAATAGTTCCTCTTTTCAACGTCATATCCGGAAACGACGCTCAA
CAGATTTTGAGTTTGACCCACATTCGAACCTTTATCATTTTACCCTGCTTTAATAAAGCCACAATGTGCTGCTT
ATGGAAGAGCCTTAGATTTAAGCCTCAACAGTATTTTCTTATTGGGCCAAACCAATTTGAAATCTTCTGACA
TTGCCTGTTTAAATCTGTCTGCAATAGCAATGCTCAAGTGTAAAGTGAAGCTGAATTTTCAGCCATTCTCATG
TCAAATATTTGGATTTGACAAACAATAGACTAGACTTTGATAATGCTAGTGTCTTACTGAATTTGTCGACTTGG
AAGTTCTAGATCTCAGCTATAATTCACACTATTTTCAAGATAGCAGGCGTAACACATCATCTAGAATTTATTCAA
ATTTCACAAATCTAAAGTTTTTAACTTGAGCCACAACAACATTTTATCTTTAAGAGATAAGTATAACCTGGAAA
GCAAGTCCCTGGTGAATTAGTTTTTCAAGTGGCAATCGCCTTGACATTTTGTGGAATGATGATGACAACAGGTATA
TCTCCATTTTCAAAGGTCTCAAGAATCTGACACGCTGATTTTATCCCTTAATAGGCTGAAGCAGCATCCCAAATG
AAGCATTCCTTAATTTGCCAGCGAGTCTCACTGAACCTACATATAAATGATAATATGTTAAAGTTTTTAACTGGA
CATTACTCCAGCAGTTTCTCGTCTCGAGTTGCTTGACTTACGTGGAAACAACTACTCTTTTAACTGATAGCC
TATCTGACTTTACATCTTCCCTTCGGACACTGCTGCTGAGTCATAACAGGATTTCCCACTACCTCTGGCTTTC
TTTCTGAAGTCAGTAGTCTGAAGCACCTCGATTTAAGTTCCAATCTGCTAAAAACAATCAACAAATCCGCACTTG
AAACTAAGACCACCACCAAAATTTATCTATGTTGGAACCTACCGGAAACCCCTTTGAATGCACCTGTGACATTGGAG
AATTCGAAGATGGATGGATGAACATCTGAATGTCAAAATTTCCAGACTGGTAGATGTCAATTTGTGCCAGTCTTG
GGGATCAAAGAGGGAAGAGTATTGTGAGTCTGGAGCTAACAACTTGTGTTTCAGATGTCACTGCAGTGATATTAT
TTTTCTTTCAGTTCTTTATCACCACCATGTTTATGTTGGCTGCCCTGGCTCACCATTGTTTTACTGGGATGTTT
GGTTTATATATAATGTGTGTTTAGCTAAGGTAAAGGCTACAGGTCTCTTCCACATCCCAAACCTTTCTATGATG
CTTACATTTCTTATGACACCAAGATGCCTCTGTTACTGACTGGGTGATAAATGAGCTGCGCTACCACTTGAAG
AGAGCCGAGACAAAAACGTTCTCCTTTGTCTAGAGGAGAGGGATTGGGACCCGGGATTGGCCATCATCGACAACC
TCATGCAGAGCATCAACCAAGCAAGAAAACAGTATTTGTTTAAACCAAAAAATATGCAAAAAGCTGGAACCTTA
AAACAGCTTTTTTACTTGGCTTTGCGAGGCTAATGGATGAGAACATGGATGTGATTATATTTATCTCTGCTGGAGC
CAGTGTACAGCATTCTCAGTATTTGAGGCTACGGCAGCGGATCTGTAAGAGCTCCATCTCCAGTGGCTGACA
ACCCGAAGGCAGAAGGCTTGTGTTTGGCAAACCTCTGAGAAATGTGGTCTTGACTGAAAAATGATTCACGGTATAACA
ATATGTATGTGATTCCATTAAGCAATACTAACTGACGTTAAGTTCATGATTTTCGCGCCATAATAAAGATGCAAG
GAATGACATTTCTGTATTAGTTATCTATTGCTATGTAACAAATTTATCCCAAACCTTAGTGGTTTTAAACAAACACA
TTTGTGCTGGCCACAGTTTTTGGGGTCAAGAGTCCAGGCCAGCATAACTGGGTCTCTGCTCAGGGTGTCTCAG
AGGCTGCAATGTAGGTGTTTACCAGAGACATAGGCATCACTGGGGTCACACTCATGTGGTTGTTTTCTGGATTCA
ATTCCTCTGGGCTATTGGCCAAAGGCTATACCTCATGTAAGCCATGCGAGCTCTCCCAAGGCAGCTTGCTTC
ATCAGAGCTAGCAAAAAGAGAGGTTGCTAGCAAGATGAAGTCACAATCTTTTGTAAATCGAATCAAAAAGTGAT
ATCTCATCACTTTGGCCATATTCTATTTGTTAGAAGTAAACACAGGTCCCAACAGCTCCATGGGAGTGACCACC
TCAGTCCAGGAAAAACAGCTGAAGACCAAGATGGTGAAGTCTGATTGCTTCAGTTGGTTCATCAACTATTTTCCCT
TGACTGCTGCTCTGGGATGGCCTGCTATCTTGATGATAGATTGTGAATATCAGGAGGCAGGATCACTGTGGACC
ATCTTAGCAGTTGACCTAACACATCTTCTTTTCAATATCTAAGAACTTTTGGCACTGTGACTAATGGTCTTAATA
TTAAGCTGTTGTTTATATTTATCATATATCTATGGCTACATGGTTATATTTATGCTGTGGTTGCGTTTCGGTTTTAT
TTACAGTTGCTTTTACAAATATTTGCTGTAACATTTGACTTCTAAGGTTTAGATGCCATTTAAGAACTGAGATGG
ATAGCTTTTAAAGCATCTTTTACTTCTTACCATTTTAAAGTATGCAGCTAAATTCGAAGCTTTTGGTCTATA
TTGTTAATTGCCATTGCTGTAAATCTTAAATGAATGAATAAAATGTTTCATTTTACAAAAA

FIGURE 211

MENMFLQSSMLTCIFLLISGSCELCAEENFSRSYPCDEKKQND SVIAECSNRRLQEVPTVG
KYVTELDLSDNFITHITNESFQGLQNLTKINLNHNPNVQH QNGNPGIQSNGLNITDGAFLNL
KNLRELLLEDNQLPQIPSGLPESLTEL SLIQNNIYNITKEGISRLINLKNLYLAWNCYFNKV
CEKTNIEDGVFETLTNLELLSLSFNSLSHVPPKLPSSLRKLFLSNTQIKYISEEDFKGLINL
TLLDLSGNCPRCFNAPFPCVPCDGGASINIDRFAFQNL TQLRYLNLSSLSLRKINA AWFKNM
PHLKVLDFEFNYLVGEIVSGAFLTMLPRLEILDLSFN YIKGSYPQHINISRNF SKLLSLRAL
HLRGYVFQELREDDFQPLMQLPNLSTINLGINF IKQIDFKLFQNF SNLEIIYLSENRI SPLV
KDTRQSYANSSSFQRHIRKRRSTDFEFDPHSNFYHFTRPLIKPQCAAYGKALDLSLNSIFFI
GPNQFENLPDIACLNL SANSNAQVLSGTEFSAIPHVKYLDLTNNRLDFDNASALTELS DLEV
LDLSYN SHYFRIAGVTHHLEFIQNFTNLKVLNLSHNNIYTLTDKYNLESKSLVELVFSGNRL
DILWNDDDNRYISIFKGLKNLTRLDLSLNRLKHIPNEAFLNLPASLTE LHINDNMLKFFNWT
LLQQFPRLELLDLRGNKLLFLTDSLSDFTSSLRTLLLSHNRISHLPSGFLSEVSSLKHL DLS
SNLLKTINKSALETKTTTTKLSMLELHG NPFECTCDIGDFRRWMDEHLNVKIPRLVDVICASP
GDQRGKSIVSLELTTCVSDVTAVILFFFTFFITTMVMLAALAHHLFYWDVWFIYNVCLAKVK
GYRSLSTSQTIFYDAYISYDTKDASVTDWVINELRYHLEESRDKNVLLCLEERDWD PGLAIID
NMQSINQSKKT VFLTKKYAKSWNFKTAFYLALQRLMDENMDVIIIFILLEPVLQHSQYLRL
RQRICKSSILQWPDNPKAEGLFWQTLRNVVLTENDSRYNMYVDSIKQY

Signal sequence:

amino acids 1-26

Transmembrane domain:

amino acids 826-848

FIGURE 212

CCAGGTCCAACCTGCACCTCGGTTCTATCGATTGAATTCCTCGGGGATCCTCTAGAGATCCCT
CGACCTCGACCCACGCGTCCGCCAAGCTGGCCCTGCACGGCTGCAAGGGAGGCTCCTGTGGA
CAGGCCAGGCAGGTGGGCCTCAGGAGGTGCCTCCAGGCGGCCAGTGGGCCTGAGGCCCCAGC
AAGGGCTAGGGTCCATCTCCAGTCCCAGGACACAGCAGCGGCCACCATGGCCACGCCTGGGC
TCCAGCAGCATCAGCAGCCCCCAGGACCGGGGAGGCACAGGTGGCCCCCACCACCCGGAGGA
GCAGCTCCTGCCCCCTGTCCGGGGGATGACTGATTCTCCTCCGCCAGGCCACCCAGAGGAGAA
GGCCACCCCGCCTGGAGGCACAGGCCATGAGGGGCTCTCAGGAGGTGCTGCTGATGTGGCTT
CTGGTGTTTGGCAGTGGGCGGCACAGAGCACGCCTACCGGCCCGGCCGTAGGGTGTGTGCTGT
CCGGGCTCACGGGGACCTGTCTCCGAGTCGTTTCGTGCAGCGTGTGTACCAGCCCTTCCTCA
CCACCTGCGACGGGCACCGGGCCTGCAGCACCTACCGAACCATCTATAGGACCGCCTACCGC
CGCAGCCCTGGGCTGGCCCCCTGCCAGGCCTCGCTACGCGTGCTGCCCCGGCTGGAAGAGGAC
CAGCGGGCTTCCTGGGGCCTGTGGAGCAGCAATATGCCAGCCGCCATGCCGGAACGGAGGGA
GCTGTGTCCAGCCTGGCCGCTGCCGCTGCCCTGCAGGATGGCGGGGTGACACTTGCCAGTCA
GATGTGGATGAATGCAGTGCTAGGAGGGGCGGCTGTCCCCAGCGCTGCATCAACACCGCCGG
CAGTTACTGGTGCCAGTGTTGGGAGGGGCACAGCCTGTCTGCAGACGGTACACTCTGTGTGC
CCAAGGGAGGGCCCCCAGGGTGGCCCCCAACCCGACAGGAGTGGACAGTGCAATGAAGGAA
GAAGTGCAGAGGCTGCAGTCCAGGGTGGACCTGCTGGAGGAGAAGCTGCAGCTGGTGCTGGC
CCCACTGCACAGCCTGGCCTCGCAGGCACTGGAGCATGGGCTCCCGGACCCCGGCAGCCTCC
TGGTGCACCTCCTTCCAGCAGCTCGGCCGCATCGACTCCCTGAGCGAGCAGATTTCTTCTCCTG
GAGGAGCAGCTGGGGTCTGCTCCTGCAAGAAAGACTCGTGACTGCCCAGCGCCCCAGGCTG
GACTGAGCCCCCTCACGCCGCCCTGCAGCCCCCATGCCCTGCCCAACATGCTGGGGGTCCAG
AAGCCACCTCGGGGTGACTGAGCGGAAGGCCAGGCAGGGCCTTCCTCCTCTTCTCCTCCCC
TTCCTCGGGAGGCTCCCCAGACCCTGGCATGGGATGGGCTGGGATCTTCTCTGTGAATCCAC
CCCTGGCTACCCCCACCCTGGCTACCCCAACGGCATCCCAAGGCCAGGTGGGCCCTCAGCTG
AGGGAAGGTACGAGCTCCCTGCTGGAGCCTGGGACCCATGGCACAGGCCAGGCAGCCCGGAG
GCTGGGTGGGGCCTCAGTGGGGGCTGCTGCCTGACCCCCAGCACAATAAAAATGAAACGTGA
AAGGGCGGCCGCGACTCTAGAGT
CGACCTGCAGAAGCTTGGCCGCCATGGCCCAACTTGTTTATTGCAGCTTATAATGGTTACAAAT

FIGURE 213

MRGSQEVLLMWLLVLAVGGTEHAYRPGRRVCAVRAHGDVSESFVQRVYQPFLTTCDGHRAC
STYRTIYRTAYRRSPGLAPARPRYACCPGWKRTSGLPGACGAAICQPPCRNGGSCVQPGRCR
CPAGWRGDTQCQSDVDECSARRGGCPQRCINTAGSYWCQCWEGHSLSADGTL CVPKGGPPRVA
PNPTGVDSAMKEEVQRLQSRVDLLEEKQLVLAPLHSLASQALEHGLPDPGSLLVHSFQQLG
RIDSLSEQISFLEEQLGSCSCKKDS

Signal sequence:

1-19

Year	Age	Sex	Occupation	Education	Income	Health	Family	Community	Environment	Policy	Program	Impact	Outcome	Conclusion
1990	25	Male	Teacher	High School	\$15,000	Good	2 Children	Urban	Good	Good	Good	Good	Good	Good
1991	26	Female	Nurse	College	\$20,000	Good	1 Child	Urban	Good	Good	Good	Good	Good	Good
1992	27	Male	Engineer	College	\$25,000	Good	2 Children	Urban	Good	Good	Good	Good	Good	Good
1993	28	Female	Doctor	College	\$30,000	Good	1 Child	Urban	Good	Good	Good	Good	Good	Good
1994	29	Male	Lawyer	College	\$35,000	Good	2 Children	Urban	Good	Good	Good	Good	Good	Good
1995	30	Female	Manager	College	\$40,000	Good	1 Child	Urban	Good	Good	Good	Good	Good	Good
1996	31	Male	Scientist	College	\$45,000	Good	2 Children	Urban	Good	Good	Good	Good	Good	Good
1997	32	Female	Artist	College	\$50,000	Good	1 Child	Urban	Good	Good	Good	Good	Good	Good
1998	33	Male	Writer	College	\$55,000	Good	2 Children	Urban	Good	Good	Good	Good	Good	Good
1999	34	Female	Designer	College	\$60,000	Good	1 Child	Urban	Good	Good	Good	Good	Good	Good
2000	35	Male	Developer	College	\$65,000	Good	2 Children	Urban	Good	Good	Good	Good	Good	Good
2001	36	Female	Analyst	College	\$70,000	Good	1 Child	Urban	Good	Good	Good	Good	Good	Good
2002	37	Male	Consultant	College	\$75,000	Good	2 Children	Urban	Good	Good	Good	Good	Good	Good
2003	38	Female	Researcher	College	\$80,000	Good	1 Child	Urban	Good	Good	Good	Good	Good	Good
2004	39	Male	Executive	College	\$85,000	Good	2 Children	Urban	Good	Good	Good	Good	Good	Good
2005	40	Female	Director	College	\$90,000	Good	1 Child	Urban	Good	Good	Good	Good	Good	Good
2006	41	Male	VP	College	\$95,000	Good	2 Children	Urban	Good	Good	Good	Good	Good	Good
2007	42	Female	CEO	College	\$100,000	Good	1 Child	Urban	Good	Good	Good	Good	Good	Good
2008	43	Male	Founder	College	\$105,000	Good	2 Children	Urban	Good	Good	Good	Good	Good	Good
2009	44	Female	Partner	College	\$110,000	Good	1 Child	Urban	Good	Good	Good	Good	Good	Good
2010	45	Male	Investor	College	\$115,000	Good	2 Children	Urban	Good	Good	Good	Good	Good	Good
2011	46	Female	Advisor	College	\$120,000	Good	1 Child	Urban	Good	Good	Good	Good	Good	Good
2012	47	Male	Coach	College	\$125,000	Good	2 Children	Urban	Good	Good	Good	Good	Good	Good
2013	48	Female	Trainer	College	\$130,000	Good	1 Child	Urban	Good	Good	Good	Good	Good	Good
2014	49	Male	Manager	College	\$135,000	Good	2 Children	Urban	Good	Good	Good	Good	Good	Good
2015	50	Female	Director	College	\$140,000	Good	1 Child	Urban	Good	Good	Good	Good	Good	Good
2016	51	Male	VP	College	\$145,000	Good	2 Children	Urban	Good	Good	Good	Good	Good	Good
2017	52	Female	CEO	College	\$150,000	Good	1 Child	Urban	Good	Good	Good	Good	Good	Good
2018	53	Male	Founder	College	\$155,000	Good	2 Children	Urban	Good	Good	Good	Good	Good	Good
2019	54	Female	Partner	College	\$160,000	Good	1 Child	Urban	Good	Good	Good	Good	Good	Good
2020	55	Male	Investor	College	\$165,000	Good	2 Children	Urban	Good	Good	Good	Good	Good	Good
2021	56	Female	Advisor	College	\$170,000	Good	1 Child	Urban	Good	Good	Good	Good	Good	Good
2022	57	Male	Coach	College										

GCCAGGCAGGTGGGCTCAGGAGGTGCCTCCAGGCGGCCAGTGGGCCTGAGGCCCCAGCAAG
GGCTAGGGTCCATCTCCAGTCCCAGGACACAGCAGCGGCCACCATGGCCACGCCTGGGCTCC
AGCAGCATCAGAGCAGCCCCCTGTGGTTGGCAGCAAAGTTCAGCTTGGCTGGGCCCCGCTGTGA
GGGGCTTCGCGCTACGCCCTGCGGTGTCCCGAGGGCTGAGGTCTCCTCATCTTCTCCCTAGC
AGTGGATGAGCAACCCAACGGGGGCCCCGGGGAGGGGAAGTGGCCCCGAGGGAGAGGAACCC
AAAGCCACATCTGTAGCCAGGATGAGCAGTGTGAATCCAGGCAGCCCCCAGGACCGGGGAGG
CACAGGTGGCCCCCACCACCCGGAGGAGCAGCTCCTGCCCCCTGTCCGGGGGATGACTGATTCT
TCCTCCGCCAGGCCACCCAGAGGAGAAGGCCACCCCGCCTGGAGGCACAGGCCATGAGGGGC
TCTCAGGAGGTGCTGCTGATGTGGCTTCTGGTGTTGGCAGTGGGCGGCACAGAGCACGCCTA
CCGGCCCCGGCCGTAGGGTGTGTGCTGTCCGGGCTCACGGGGACCCTGTCTCCGAGTCGTTTCG
TGCAGCGTGTGTACCAGCCCTTCCTCACCACCTGCGACGGGCACCGGGCCTGCAGCACCTAC
CGAACCATCTATAGGACCGCCTACCGCCGCAGCCCTGGGCTGGCCCCCTGCCAGGCCTCGCTA
CGCGTGCTGCCCCGGCTGGAAGAGGACCAGCGGGCTTCCTGGGGCCTGTGGAGCAGCAATAT
GCCAGCCGCCATGCCGGAACGGAGGGAGCTGTGTCCAGCCTGGCCGCTGCCGCTGCCCTGCA
GGATGGCGGGGTGACACTTGCCAGTCAGATGTGGATGAATGCAGTGCTAGGAGGGGCGGCTG
TCCCCAGCGCTGCATCAACACCGCCGGCAGTTACTGGTGCCAGTGTTGGGAGGGGCACAGCC
TGTCTGCAGACGGTACACTCTGTGTGCCCAAGGGAGGGCCCCCAGGGTGGCCCCCAACCCG
ACAGGAGTGGACAGTGCAATGAAGGAAGAAGTGCAGAGGCTGCAGTCCAGGGTGGACCTGCT
GGAGGAGAAGCTGCAGCTGGTGCTGGCCCCACTGCACAGCCTGGCCTCGCAGGCACTGGAGC
ATGGGCTCCCGGACCCCGGCAGCCTCCTGGTGCCTCCTTCCAGCAGCTCGGCCGCATCGAC
TCCCTGAGCGAGCAGATTTCTTCTCCTGGAGGAGCAGCTGGGGTCTGCTCCTGCAAGAAAGA
CTCGTGACTGCCCAGCGCTCCAGGCTGGACTGAGCCCCCTCACGCCGCCCTGCAGCCCCCATG
CCCCCTGCCCAACATGCTGGGGGTCCAGAAGCCACCTCGGGGTGACTGAGCGGAAGGCCAGGC
AGGGCCTTCTCCTCTTCTCCTCCTCCCCTTCTCGGGAGGCTCCCCAGACCCTGGCATGGGAT
GGGCTGGGATCTTCTCTGTGAATCCACCCCTGGCTACCCCCACCCTGGCTACCCCAACGGCA
TCCCAAGGCCAGGTGGACCCCTCAGCTGAGGGAAGGTACGAGCTCCCTGCTGGAGCCTGGGAC
CCATGGCACAGGCCAGGCAGCCCGGAGGCTGGGTGGGGCCTCAGTGGGGGCTGCTGCCTGAC
CCCCAGCACAATAAAAAATGAAACGTG

FIGURE 215

MRGSQEVLLMWLLVLAVGGTEHAYRPGRRVCAVRAHGDPVSESFVQRVYQPFLTTCDGHRAC
STYRTIYRTAYRRSPGLAPARPRYACCPGWKRTSGLPGACGAAICQPPCRNGGSCVQPGRCR
CPAGWRGDTQCSDVDECSARRGGCPQRCINTAGSYWCQCWEGHSLSadgtLCVpKGGPPRVA
PNPTGVDSAMKEEVQRLQSRVDLLEEKQLVLAPLHSLASQALEHGLPDPGSLLVHSFQQLG
RIDSLSEQISFLEEQLGSCSCKKDS

Signal sequence:

1-19

bioRxiv preprint doi: <https://doi.org/10.1101/2020.03.10.288888>; this version posted March 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 216

CCCACGCGTCCGAAGCTGGCCCTGCACGGCTGCAAGGGAGGCTCCTGTGGACAGGCCAGGCA
GGTGGGCCTCAGGAGGTGCCTCCAGGCGGCCAGTGGGCCTGAGGCCCCAGCAAGGGCTAGGG
TCCATCTCCAGTCCCAGGACACAGCAGCGGCCACCATGGCCACGCCTGGGCTCCAGCAGCAT
CAGCAGCCCCCAGGACCGGGGAGGCACAGGTGGCCCCCACCACCCGGAGGAGCAGCTCCTGC
CCCTGTCCGGGGGATGACTGATTCTCCTCCGCCAGGCCACCCAGAGGAGAAGGCCACCCCGC
CTGGAGGCACAGGCCATGAGGGGCTCTCAGGAGGTGCTGCTGATGTGGCTTCTGGTGTGGC
AGTGGGCGGCACAGAGCACGCCTACCGGCCCGGCCGTAGGGTGTGTGCTGTCCGGGCTCACG
GGGACCCTGTCTCCGAGTCGTTTCGTGCAGCGTGTGTACCAGCCCTTCCTCACCACCTGCGAC
GGGCACCGGGCCTGCAGCACCTACCGAACCATCTATAGGACCGCCTACCGCCGCAGCCCTGG
GCTGGCCCCCTGCCAGGCCTCGCTACGCGTGCTGCCCCGGCTGGAAGAGGACCAGCGGGCTTC
CTGGGGCCTGTGGAGCAGCAATATGCCAGCCGCCATGCCGGAACGGAGGGAGCTGTGTCCAG
CCTGGCCGCTGCCGCTGCCCTGCAGGATGGCGGGGTGACACTTGCCAGTCAGATGTGGATGA
ATGCAGTGCTAGGAGGGGCGGCTGTCCCCAGCGCTGCGTCAACACCGCCGGCAGTTACTGGT
GCCAGTGTGGGAGGGGCACAGCCTGTCTGCAGACGGTACACTCTGTGTGCCCAAGGGAGGG
CCCCCAGGGTGGCCCCCAACCCGACAGGAGTGGACAGTGCAATGAAGGAAGAAGTGCAGAG
GCTGCAGTCCAGGGTGGACCTGCTGGAGGAGAAGCTGCAGCTGGTGCTGGCCCCACTGCACA
GCCTGGCCTCGCAGGCACTGGAGCATGGGCTCCCGGACCCCGGCAGCCTCCTGGTGCCTCC
TTCCAGCAGCTCGGCCGCATCGACTCCCTGAGCGAGCAGATTTCTTCCTGGAGGAGCAGCT
GGGGTCCTGCTCCTGCAAGAAAGACTCGTGACTGCCCAGCGCCCCAGGCTGGACTGAGCCCC
TCACGCCGCCCTGCAGCCCCCATGCCCTGCCCAACATGCTGGGGGTCCAGAAGCCACCTCG
GGGTGACTGAGCGGAAGGCCAGGCAGGGCCTTCCTCCTCTTCCTCCTCCCCTTCCTCGGGAG
GCTCCCCAGACCCTGGCATGGGATGGGCTGGGATCTTCTCTGTGAATCCACCCCTGGCTACC
CCCACCCTGGCTACCCAACGGCATCCCAAGGCCAGGTGGGCCCTCAGCTGAGGGAAGGTAC
GAGCTCCCTGCTGGAGCCTGGGACCCATGGCACAGGCCAGGCAGCCCGAGGCTGGGTGGGG
CCTCAGTGGGGGCTGCTGCCTGACCCCCAGCACATAAAAATGAAACGTG

FIGURE 217

MRGSQEVLLMWLLVLAVGGTEHAYRPGRRVCAVRAHGDPVSESFVQRVYQPFLTTCDGHRAC
STYRTIYRTAYRRSPGLAPARPRYACCPGWKRTSGLPGACGAAICQPPCRNGGSCVQPGRCR
CPAGWRGDTQCQSDVDECSARRGGCPQRCVNTAGSYWCQCWEGHSLSADGTLCVPKGPPRVA
PNPTGVDSAMKEEVQRLQSRVDLLEEKQLVLAPLHSLASQALEHGLPDPGSLLVHSFQQLG
RIDSLSEQISFLEEQLGSCSCKKDS

Signal sequence:

1-19

Variable	Mean	Standard deviation	Minimum	Maximum
Age	35.2	12.5	18	65
Gender	0.52	0.50	0	1
Marital status	0.68	0.48	0	1
Education	12.5	2.1	9	16
Income	15.2	8.5	5	35
Health status	0.75	0.43	0	1
Employment status	0.82	0.38	0	1
Home ownership	0.91	0.29	0	1
Vehicle ownership	0.78	0.41	0	1
Life satisfaction	4.2	1.8	1	7
Subjective health	3.8	1.5	1	6
Life expectancy	78.5	5.2	65	90
Healthcare expenditure	12.1	6.3	3	28
Health insurance coverage	0.95	0.23	0	1
Physical activity	0.65	0.47	0	1
Smoking status	0.25	0.43	0	1
Alcohol consumption	0.35	0.48	0	1
Dietary habits	0.55	0.50	0	1
Stress levels	4.5	2.0	1	7
Social support	5.2	1.9	1	7
Community engagement	3.5	1.7	1	6
Environmental quality	4.8	1.6	1	6
Access to green spaces	0.72	0.45	0	1
Public transportation	0.85	0.35	0	1
Urbanization level	0.98	0.15	0	1
Government spending on health	18.5	7.2	8	32
Healthcare quality index	6.5	1.2	5	8
Healthcare accessibility	0.88	0.32	0	1
Healthcare equity	0.75	0.42	0	1
Healthcare innovation	0.62	0.49	0	1
Healthcare sustainability	0.58	0.51	0	1
Healthcare governance	0.70	0.46	0	1
Healthcare transparency	0.65	0.48	0	1
Healthcare accountability	0.60	0.49	0	1
Healthcare effectiveness	0.78	0.41	0	1
Healthcare efficiency	0.72	0.45	0	1
Healthcare safety	0.85	0.35	0	1
Healthcare quality of care	0.82	0.38	0	1
Healthcare patient satisfaction	0.75	0.42	0	1
Healthcare provider satisfaction	0.68	0.48	0	1
Healthcare system performance	0.70	0.46	0	1
Healthcare system resilience	0.65	0.48	0	1
Healthcare system sustainability	0.60	0.49	0	1
Healthcare system equity	0.55	0.50	0	1
Healthcare system transparency	0.50	0.51	0	1
Healthcare system accountability	0.45	0.52	0	1
Healthcare system effectiveness	0.40	0.53	0	1
Healthcare system efficiency	0.35	0.54	0	1
Healthcare system safety	0.30	0.55	0	1
Healthcare system quality of care	0.25	0.56	0	1
Healthcare system patient satisfaction	0.20	0.57	0	1
Healthcare system provider satisfaction	0.15	0.58	0	1
Healthcare system performance	0.10	0.59	0	1
Healthcare system resilience	0.05	0.60	0	1
Healthcare system sustainability	0.00	0.61	0	1
Healthcare system equity	0.00	0.62	0	1
Healthcare system transparency	0.00	0.63	0	1
Healthcare system accountability	0.00	0.64	0	1
Healthcare system effectiveness	0.00	0.65	0	1
Healthcare system efficiency	0.00	0.66	0	1
Healthcare system safety	0.00	0.67	0	1
Healthcare system quality of care	0.00	0.68	0	1
Healthcare system patient satisfaction	0.00	0.69	0	1
Healthcare system provider satisfaction	0.00	0.70	0	1
Healthcare system performance	0.00	0.71	0	1
Healthcare system resilience	0.00	0.72	0	1
Healthcare system sustainability	0.00	0.73	0	1
Healthcare system equity	0.00	0.74	0	1
Healthcare system transparency	0.00	0.75	0	1
Healthcare system accountability	0.00	0.76	0	1
Healthcare system effectiveness	0.00	0.77	0	1
Healthcare system efficiency	0.00	0.78	0	1
Healthcare system safety	0.00	0.79	0	1
Healthcare system quality of care	0.00	0.80	0	1
Healthcare system patient satisfaction	0.00	0.81	0	1
Healthcare system provider satisfaction	0.00	0.82	0	1
Healthcare system performance	0.00	0.83	0	1
Healthcare system resilience	0.00	0.84	0	1
Healthcare system sustainability	0.00	0.85	0	1
Healthcare system equity	0.00	0.86	0	1
Healthcare				

Variable	Mean	Standard deviation	Minimum	Maximum
Age	35.2	12.5	18	65
Gender	0.52	0.50	0	1
Marital status	0.68	0.48	0	1
Education	12.5	2.1	9	16
Income	15.2	8.5	5	35
Health status	0.75	0.43	0	1
Employment status	0.82	0.38	0	1
Home ownership	0.91	0.29	0	1
Vehicle ownership	0.78	0.41	0	1
Life satisfaction	4.2	1.8	1	7
Subjective health	3.8	1.5	1	6
Life expectancy	78.5	5.2	65	90
Healthcare expenditure	12.1	4.5	5	25
Health insurance coverage	0.95	0.22	0	1
Physical activity	0.65	0.48	0	1
Smoking status	0.25	0.43	0	1
Alcohol consumption	0.35	0.48	0	1
Dietary habits	0.55	0.50	0	1
Stress levels	4.5	1.2	1	7
Social support	5.2	1.5	1	7
Community involvement	0.72	0.45	0	1
Environmental quality	3.5	1.2	1	6
Access to healthcare	0.88	0.32	0	1
Healthcare quality	4.8	1.5	1	7
Healthcare accessibility	0.92	0.28	0	1
Healthcare affordability	0.85	0.35	0	1
Healthcare quality of life	4.5	1.2	1	7
Healthcare satisfaction	4.2	1.5	1	7
Healthcare equity	0.78	0.41	0	1
Healthcare transparency	0.82	0.38	0	1
Healthcare accountability	0.88	0.32	0	1
Healthcare effectiveness	0.92	0.28	0	1
Healthcare efficiency	0.85	0.35	0	1
Healthcare safety	0.95	0.22	0	1
Healthcare quality of care	4.8	1.5	1	7
Healthcare patient satisfaction	4.2	1.5	1	7
Healthcare provider satisfaction	4.5	1.2	1	7
Healthcare system performance	4.8	1.5	1	7
Healthcare system quality	4.5	1.2	1	7
Healthcare system effectiveness	4.8	1.5	1	7
Healthcare system efficiency	4.5	1.2	1	7
Healthcare system safety	4.8	1.5	1	7
Healthcare system quality of care	4.8	1.5	1	7
Healthcare system patient satisfaction	4.2	1.5	1	7
Healthcare system provider satisfaction	4.5	1.2	1	7
Healthcare system performance	4.8	1.5	1	7
Healthcare system quality	4.5	1.2	1	7
Healthcare system effectiveness	4.8	1.5	1	7
Healthcare system efficiency	4.5	1.2	1	7
Healthcare system safety	4.8	1.5	1	7
Healthcare system quality of care	4.8	1.5	1	7
Healthcare system patient satisfaction	4.2	1.5	1	7
Healthcare system provider satisfaction	4.5	1.2	1	7
Healthcare system performance	4.8	1.5	1	7
Healthcare system quality	4.5	1.2	1	7
Healthcare system effectiveness	4.8	1.5	1	7
Healthcare system efficiency	4.5	1.2	1	7
Healthcare system safety	4.8	1.5	1	7
Healthcare system quality of care	4.8	1.5	1	7
Healthcare system patient satisfaction	4.2	1.5	1	7
Healthcare system provider satisfaction	4.5	1.2	1	7
Healthcare system performance	4.8	1.5	1	7
Healthcare system quality	4.5	1.2	1	7
Healthcare system effectiveness	4.8	1.5	1	7
Healthcare system efficiency	4.5	1.2	1	7
Healthcare system safety	4.8	1.5	1	7
Healthcare system quality of care	4.8	1.5	1	7
Healthcare system patient satisfaction	4.2	1.5	1	7
Healthcare system provider satisfaction	4.5	1.2	1	7
Healthcare system performance	4.8	1.5	1	7
Healthcare system quality	4.5	1.2	1	7
Healthcare system effectiveness	4.8	1.5	1	7
Healthcare system efficiency	4.5	1.2	1	7
Healthcare system safety	4.8	1.5	1	7
Healthcare system quality of care	4.8	1.5	1	7
Healthcare system patient satisfaction	4.2	1.5	1	7
Healthcare system provider satisfaction	4.5	1.2	1	7
Healthcare system performance	4.8	1.5	1	7
Healthcare system quality	4.5	1.2	1	7
Healthcare system effectiveness	4.8	1.5	1	7
Healthcare system efficiency	4.5	1.2	1	7
Healthcare system safety	4.8	1.5	1	7

[illegible]

Putative transmembrane domains:

amino acids 36-55 (type II TM), 65-84, 188-208, 229-245

FIGURE 220

AAAACCCTGTATTTTTTACAATGCAAATAGACAATNANCCTGGAGGTCTTTGAATTAGGTAT
TATAGGGATGGTGGGGTTGATTTTTNTTCCTGGAGGCTTTTGGCTTTGGACTCTCNCTTTCT
CCCACAGAGCNCTTCGACCATCACTGCCCCTGGGTGGGGAATTGTGTTGGAAAGAGGAACTA
CCGCTANTTCTACCTCTTCATCCTTTNTCTCTCCNCCTCACAATCTATGTCTTCGCCTTCA
ACATCGT

FIGURE 221

GTTGTGTCCTTCAGCAAAACAGTGGATTTAAATCTCCTTGACACAAGCTTGAGAGCAACACAA
TCTATCAGGAAAGAAAGAAAGAAAAAACCGAACCTGACAAAAAGAAGAAAAAGAAGA
AAAAAATCATGAAAACCATCCAGCCAAAAATGCACAATTCTATCTCTTGGGCAATCTTCAC
GGGGCTGGCTGCTCTGTGTCTCTTCCAAGGAGTGCCCGTGCGCAGCGGAGATGCCACCTTCC
CCAAAGCTATGGACAACGTGACGGTCCGGCAGGGGGAGAGCGCCACCCTCAGGTGCACTATT
GACAACCGGGTCACCCGGGTGGCCTGGCTAAACCGCAGCACCATCCTCTATGCTGGGAATGA
CAAGTGGTGCCTGGATCCTCGCGTGGTCCTTCTGAGCAACACCCAAACGCAGTACAGCATCG
AGATCCAGAACGTGGATGTGTATGACGAGGGCCCTTACACCTGCTCGGTGCAGACAGACAAC
CACCCAAAGACCTCTAGGGTCCACCTCATTGTGCAAGTATCTCCCAAATTGTAGAGATTTCT
TTCAGATATCTCCATTAATGAAGGGAACAATATTAGCCTCACCTGCATAGCAACTGGTAGAC
CAGAGCCTACGGTTACTTGGAGACACATCTCTCCCAAAGCGGTTGGCTTTGTGAGTGAAGAC
GAATACTTGGAATTCAGGGCATCACCCGGGAGCAGTCAGGGGACTACGAGTGCAGTGCCTC
CAATGACGTGGCCGCGCCCGTGGTACGGAGAGTAAAGGTACCCGTGAACCTATCCACCATA
TTTCAGAAGCCAAGGGTACAGGTGTCCCGTGGGACAAAAGGGGACACTGCAGTGTGAAGCC
TCAGCAGTCCCCTCAGCAGAATTCAGTGGTACAAGGATGACAAAAGACTGATTGAAGGAAA
GAAAGGGGTGAAAGTGGAAAACAGACCTTTCTCTCAAACTCATCTTCTTCAATGTCTCTG
AACATGACTATGGGAACTACACTTGCCTGGCCTCCAACAAGCTGGGCCACACCAATGCCAGC
ATCATGCTATTTGGTCCAGGCGCCGTGAGCGAGGTGAGCAACGGCACGTCGAGGAGGGCAGG
CTGCGTCTGGCTGCTGCCTCTTCTGGTCTTGACCTGCTTCTCAAATTTTTGATGTGAGTGCC
ACTTCCCCACCCGGGAAAGGCTGCCGCCACCACCACCACCAACACAACAGCAATGGCAACAC
CGACAGCAACCAATCAGATATATACAAATGAAATTAGAAGAAACACAGCCTCATGGGACAGA
AATTTGAGGGAGGGGAACAAAGAATACTTTGGGGGGAAAAGAGTTTTAAAAAAGAAATTGAA
AATTGCCTTGCAGATATTTAGGTACAATGGAGTTTTCTTTTCCCAAACGGGAAGAACACAGC
ACACCCGGCTTGGACCCACTGCAAGCTGCATCGTGCAACCTCTTTGGTGCCAGTGTGGGCAA
GGGCTCAGCCTCTCTGCCACAGAGTGCCCCACGTGGAACATTCTGGAGCTGGCCATCCCA
AATTCAATCAGTCCATAGAGACGAACAGAATGAGACCTTCCGGCCCAAGCGTGGCGCTGCGG
GCACTTTGGTAGACTGTGCCACCACGGCGTGTGTTGTGAAACGTGAAATAAAAAGAGCAAAA
AAAAA

FIGURE 222

MKTIQPKMHNSISWAIIFTGLAALCLFQGVFVRSGDATFPKAMDNVTVRQGESATLRCTIDNR
VTRVAWLNRSTILYAGNDKWCLDPRVVLLSNTQTQYSIEIQNVDVYDEGPYTCVQTDNHPK
TSRVHLIVQVSPKIVEISSDISINEGNNISLTCIATGRPEPTVTWRHISPKAVGFVSEDEYL
EIQGITREQSGDYECASNDVAAPVRRVKVTVNYPPISEAKGTGVPVGQKGTQCEASAV
PSAEFQWYKDDKRLIEGKKGVKVENRPFLSKLIFFNVSEHDYGNITCVASNKLGHTNASIML
FGPGAVSEVSNGTSRRAGCVWLLPLLVLHLLLLKF

Signal peptide:

amino acids 1-28

protein = 230000

[illegible][illegible]

FIGURE 224

ATGGCTGGTGACGGCGGGGCCGGGCAGGGGACCGGGGCCGCGGGCCGGGAGCGGGCCAGCTGCCGGGAGCCCTGA
ATCACCGCCTGGCCCCGACTCCACCA**ATGA**ACGTGCGCGCTGCAGGAGCTGGGAGCTGGCAGCAACGTGGGATTCCAG
AAGGGGACAAGACAGCTGTTAGGCTCACGCACGCAGCTGGAGCTGGTCTTAGCAGGTGCCCTCTCTACTGCTGGCT
GCACTGCTTCTGGGCTGCCTTGTGGCCCTAGGGGTCCAGTACCACAGAGACCCATCCCACAGCACCTGCCCTTACA
GAGGCCTGCATTGAGTGGCTGGAAAAATCCTGGAGTCCCTGGACCGAGGGGTGAGCCCCCTGTGAGGACTTTTAC
CAGTTCTCCTGTGGGGGCTGGATTCCGAGGAACCCCTGCCCGATGGGCGTTCTCGCTGGAAACACCTTCAACAGC
CTCTGGGACCAAAACCAGGCCATACTGAAGCACCTGCTTGAAAAACACCACCTTCAACTCCAGCAGTGAAGCTGAG
CAGAAGACACAGCGCTTCTACCTATCTTGCCTACAGGTGGAGCGCATTGAGGAGCTGGGAGCCCAGCCACTGAGA
GACCTCATTGAGAAGATTGGTGGTTGGAACATTACGGGGCCCTGGGACCAGGACAACTTTATGGAGGTGTTGAAG
GCAGTAGCAGGGACCTACAGGGCCACCCCATTTCTTACCCTCTACATCAGTGCCGACTCTAAGAGTTCCAACAGC
AATGTTATCCAGGTGGACCAGTCTGGGCTCTTTCTGCCCTCTCGGGATTACTACTTAAACAGAACTGCCAATGAG
AAAGTGCTCACTGCCTATCTGGATTACATGGAGGAACCTGGGGATGCTGCTGGGTGGGCGGGCCACCTCCACGAGG
GAGCAGATGCAGCAGGTGCTGGAGTTGGAGATACAGCTGGCCAAACATCACAGTGCCCCAGGACCAGCGGCGGCAG
GAGGAGAAGATCTACCACAAGATGAGCATTTCGGAGCTGCAGGCTCTGGCGCCCTCCATGGACTGGCTTGAGTTC
CTGTCTTTCTTGCTGTCAACATTGGAGTTGAGTGACTCTGAGCCTGTGGTGGTGTATGGGATGGATTATTTGCAG
CAGGTGTGAGAGCTCATCAACCGCACGGAACCAAGCATCCTGAACAATTACCTGATCTGGAACCTGGTGCAAAAG
ACAACCTCAAGCCTGGACCGACGCTTTGAGTCTGCACAAGAGAAGCTGCTGGAGACCCTCTATGGCACTAAGAAG
TCCTGTGTGCCGAGGTGGCAGACCTGCATCTCCAACACGGATGACGCCCTTGGCTTTGCTTTGGGGTCACTCTTC
GTGAAGGCCACGTTTGACCGGCAAAGCAAAGAAATTGCAGAGGGGATGATCAGCGAAATCCGGACCGCATTGAG
GAGGCCCTGGGACAGCTGGTTTGGATGGATGAGAAGACCCGCCAGGCAGCCAAGGAGAAAGCAGATGCCATCTAT
GATATGATTGGTTTCCCAGACTTTATCCTGGAGCCCAAAGAGCTGGATGATGTTTATGACGGGTACGAAATTTCT
GAAGATTCTTTCTTCCAAAACATGTTGAATTTGTACAACCTTCTCTGCCAAGGTTATGGCTGACCAGCTCCGCAAG
CCTCCCAGCCGAGACCAGTGGAGCATGACCCCCCAGACAGTGAATGCCTACTACCTTCCAACCTAAGAATGAGATC
GTCTTCCCCGCTGGCATCCTGCAGGCCCCCTTCTATGCCCGCAACCACCCCAAGGCCCTGAACCTTCGGTGGCATC
GGTGTGGTCACTGGGCCATGAGTTGACGCATGCCTTTGATGACCAAGGGCGCGAGTATGACAAAGAAGGGAACTG
CGGCCCTGGTGGCAGAATGAGTCCCTGGCAGCCTTCCGGAACCACACGGCCTGCATGGAGGAACAGTACAATCAA
TACCAGGTCAATGGGGAGAGGCTCAACGGCCGCCAGACGCTGGGGGAGAAACATTACTGACAACGGGGGGCTGAAG
GCTGCCTACAATGCTTACAAAGCATGGCTGAGAAAGCATGGGGAGGAGCAGCAACTGCCAGCCGTGGGGCTCACC
AACCACCAGCTCTTCTTCGTGGGATTGCCCCAGGTGTGGTGTCTCGGTCCGCACACCAGAGAGCTCTCACGAGGGG
CTGGTGACCGACCCCCACAGCCCTGCCCGCTTCCGCGTGTGGGCACTCTCTCCAACCTCCCGTGACTTCTCTGCGG
CACTTCGGCTGCCCTGTGCGCTCCCCCATGAACCCAGGGCAGCTGTGTGAGGTGTGGT**TAG**ACCTGGATCAGGGGA
GAAATGGCCAGCTGTCAACAGACCTGGGGCAGCTCTCCTGACAAAGCTGTTTGCTCTTGGGTGGGAGGAAGCAA
ATGCAAGCTGGGCTGGGTCTAGTCCCTCCCCCCCACAGGTGACATGAGTACAGACCCCTCCTCAATCACCACATTG
TGCTCTGCTTTGGGGGTGCCCTGCCCTCCAGCAGAGCCCCCACCATTCACTGTGACATCTTTCCGTGTACCCCT
GCCTGGAAGAGGTCTGGGTGGGGAGGCCAGTTCCCATAGGAAGGAGTCTGCC

FIGURE 225

MNVALQELGAGSNVGFQKGTRQLLGSRTQLELVLAGASLLLAALLLGCLVALGVQYHRDPSH
STCLTEACIRVAGKILESIDRGVSPCEDFYQFSCGGWIRRNPLPDGRSRWNTFNSLWDQNQA
ILKHLLENTTFNSSSEAEQKTQRFYLSCLQVERIEELGAQPLRDLEKIGGWNITGPWDQDN
FMEVLKAVAGTYRATPFFTVYISADSKSSNSNVIQVDQSGFLPSRDYYLNRTANEKVLTA
LDYMEELGMLLGGRPTSTREQMQQVLELEIQLANITVPQDQRRDEEKIYHKMSISELQALAP
SMDWLEFLSFLLSPLELSDSEPVVVGMDYLQQVSELINRTEPSILNNYLIWNLVQKTTSSL
DRRFESAQEKLETLYGTTKSCVPRWQTCISNTDDALGFALGSLFVKATFDRQSKEIAEGMI
SEIRTA FEEALGQLVWMDEKTRQAAKEKADAIYDMIGFPDFILEPKELDDVYDGYEISEDSF
FQNMLNLYNFSAKVMADQLRKPPSRDQWSMTPQTVNAYYLPTKNEIVFPAGILQAPFYARNH
PKALNFGGIGVVMGHELTHAFDDQGREYDKEGNLRPWWQNESLAAFRNHTACMEEQYNQYQV
NGERLNGRQTLGENITDNGGLKAAYNAYKAWLRKHGEEQQLPAVGLTNHQLFFVGFAQVWCS
VRTPESSHEGLVTDPHSPARFRVLGTLSNSRDFLRHFGCPVGSPMNPGLCEVW

Type II Transmembrane domain:

amino acids 32-57

[illegible][illegible]

FIGURE 227

GGCCGAGCGGGGGTGCTGCGCGGCGGCCGTGATGGCTGGTGACGGCGGGGCCGGGCAGGGGA
CCGGGGCCGCGGGCCGGGAGCGGGCCAGCTGCCGGGAGCCCTGAATCACCGCCTGGCCCCGAC
TCCACCATGAACGTCGCGCTGCAGGAGCTGGGAGCTGGCAGCAACGTGGGATTCCAGAAGGG
GACAAGACAGCTGTTAGGCTCACGCACGCAGCTGGAGCTGGTCTTAGCAGGTGCCTCTCTAC
TGCTGGCTGCACTGCTTCTGGGCTGCCTTGTGGCCCTAGGGGTCCAGTACCACAGAGACCCA
TCCCACAGCACCTGCCTTACAGAGGCCTGCATTTCGAGTGGCTGGAAAAATCCTGGAGTCCCT
GGACCGAGGGGTGAGCCCCTGTGAGGACTTTTACCAGTTCTCCTGTGGGGGCTGGATTTCGGA
GGAACCCCTGCCCGATGGGCGTTCTCGCTGGAACACCTTCAACAGCCTCTGGGACCAAAAC
CAGGCCATACTGAAGCACCTGCTTGAAAACACCACCTTCAACTCCAGCAGTGAAGCTGAGCA
GAAGACACAGCGCTTCTACCTATCTTGCCTACAGGTGGAGCGCATTGAGGAGCTGGGAGCCC
AGCCACTGAGAGACCTCATTGAGAAGATTGGTGGTTGGAACATTACGGGGCCCTGGGACCAG
GACAACTTTATGGAGGTGTTGAAGGCAGTAGCAGGGACCTACAGGGCCACCCCATTTCTTCAC
CGTCTACATCAGTGCCGACTCTAAGAGTTCCAACAGCAATGTTATCCAGGTGGACCAGTCTG
GGCTCTTTCTGCCCTCTCGGGATTACTACTTAAACAGAACTGCCAATGAGAAAGTAAGGAAC
ATCTTCCGAACCCCCATCCCTACCCCTGGCTGAGCTGGGCTGATCCCTGTTGACTTTTCCCT
TTGCCAAGGGTCAGAGCAGGGAAGGTGAGCCTATCCTGTACCTAGTGAACAACTGCCCCCT
CCTTTCTTTCTTCTTTTCTTCCCTCCCTCCCTTTCTTCCCCCTTTTCCTTCCTTCCTTCC
TCTTATTCTTCTAGTAGGTTTCATAGACACCTACTGTGTGCCAGGTCCAGTGGGGGAATTTCG
GAGATATAAGTTTCCGAGCCATTGCCACAGGAAGCGTTCAGTGTCGATGGGTTCATGGACCT
AGATAGGCTGATAACAAAGCTCACAGAGGGTCCTGAGGATTTCAGGAGAGACTTATGGAGCC
AGCAAAGTCTTCTGAAGAGATTGCATTTGAGCCAGGTCCTGTAG

FIGURE 228

ATGCCTACTACCTTCCAATAAGAATGAGATCGTCTTCCCCGCTGGCATCCTGCAGGCCCC
TTCTATGCCCCGAACCACCCCAAGGCCCTGAACTTCGGTGGCATCGGTGTGGTCATGGGCCA
TGAGTTGACGCATGCCTTTGATGACCAAGGGCGCGAGTATGACAAAGAAGGGAACCTGCGGC
CCTGGTGGCAGAATGAGTCCCTGGCAGCCTTCCGGAACCACACGGCCTGCATGGAGGAACAG
TACAATCAATACCAGGTCAATGGGGAGAGGCTCAACGGCCGCCAGACGCTGGGGGAGAACAT
TGCTGACAACGGGGGGCTGAAGGCTGCCTACAATGCTTACAAAGCATGGCTGAGAAAGCATG
GGGAGGAGCAGCAACTGCCAGCCGTGGGGCTCACCAACCACCAGCTCTTCTTCGTGGGATTT
GCCCAGGTGTGGTGCTCGGTCCGCACACCAGAGAGCTCTCACGAGGGGGCTGGTGACCGACCC
CCACAGCCCTGCCCCGCTTCCGCGTGCTGGGCACTCTCTCCAACCTCCCGTGACTTCCTGCGGC
ACTTCGGCTGCCCTGTTCGGCTCCCCCATGAACCCAGGGCAGCTGTGTGAGGTGTGGTAGACC
TGGATCAGGGGAGAAAATGGCCAGCTGTCAACCAGACCTGGGGCAGCTCTCCTGACAAAGCTGT
TTGCTCTTGGGTTGGGAGGAAGCAAATGCAAGCTGGGCTGGGTCTAGTCCCTCCCCCCCCACA
GGTGACATGAGTACAGACCCTCCTCAATCACCACATTGTGCCTCTGCTTTGGGGGTGCCCCCT
GCCTCCAGCAGAGCCCCCACCATTCACTGTGACATCTTTCGCTGTACCCCTGCCTGGAAGAG
GTCTGGGTGGGGAGGCCAGTTCCCATAGGAAGGAGTCTGCCTCTTCTGTCCCCAGGCTCACT
CAGCCTGGCGGCCATGGGGCTGCCGTGCCTGCCCCACTGTGACCCACAGGCCTGGGTGGTG
TACCTCCTGGACTTCTCCCCAGGCTCACTCAGTGCGCACTTAGGGGTGGACTCAGCTCTGTC
TGGCTCACCCCTCACGGGCTACCCCCACCTCACCCCTGTGCTCCTTGTGCCACTGCTCCCAGTG
CTGCTGCTGACCTTCACTGACAGCTCCTAGTGGAAGCCCCAAGGGCCTCTGAAAGCCTCCTGC
TGCCCACTGTTTCCCTGGGCTGAGAGGGGAAGTGCATATGTGTAGCGGGTACTGGTTCCCTGT
GTCTTAGGGCACAAGCCTTAGCAAATGATTGATTCTCCCTGGACAAAGCAGGAAAGCAGATA
GAGCAGGGAAAAGGAAGAACAGAGTTTATTTTACAGAAAAGAGGGTGGGAGGGTGTGGTCT
TGGCCCTTATAGGACC

1002967 102401

FIGURE 229

CCCACGCGTCCGAGCCGCCGAGAATTAGACACACTCCGGACGCGGCCAAAAGCAACCGAGA
GGAGGGGAGGCAAAAACACCGAAAAACAAAAGAGAGAAACAACACCCAACTGGGGTGG
GGGGAAGAAAGAAAGAAAAGAAACCCACCCACCCACCAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAATCCTGTGGCGCGCCGCTGGTTCCCGGAAGACTCGCCAGCACCAGGGGG
TGGGGGAGTGCGAGCTGAAAGCTGCTGGAGAGTGAGCAGCCCTAGCAGGGATGGACATGATG
CTGTTGGTGCAGGGTGCTTGTGCTCGAACCAAGTGGCTGGCGGCGGTGCTCCTCAGCCTGTG
CTGCCTGCTACCCTCCTGCCTCCCGGCTGGACAGAGTGTGGACTTCCCCCTGGGCGGCCGTGG
ACAACATGATGGTCAGAAAAGGGGACACGGCGGTGCTTAGGTGTTATTTGGAAGATGGAGCT
TCAAAGGGTGCTGGCTGAACCGGTCAAGTATTTATTTTTCGCGGAGGTGATAAGTGGTCAGT
GGATCCTCGAGTTTCAATTTCAACATTGAATAAAAGGGACTACAGCCTCCAGATACAGAATG
TAGATGTGACAGATGATGGCCATACACGTGTTCTGTTTCACTCAACATACACCCAGAACA
ATGCAGGTGCATCTAACTGTGCAAGTTCCTCCTAAGATATATGACATCTCAAATGATATGAC
CGTCAATGAAGGAACCAACGTCACTCTTACTTGTGTTGGCCACTGGGAAACCAGAGCCTTCCA
TTTCTTGGCGACACATCTCCCCATCAGCAAAACCATTGAAAAATGGACAATATTTGGACATT
TATGGAATTACAAGGGACCAGGCTGGGGAATATGAATGCAGTGCAGGAAATGCTGTGTGCTATT
CCCAGATGTGAGGAAAGTAAAAGTTGTTGTCAACTTTGCTCCTACTATTCAGGAAATTAAAT
CTGGCACCGTGACCCCCGGACGCAGTGGCCTGATAAGATGTGAAGGTGCAGGTGTGCCGCT
CCAGCCTTTGAATGGTACAAAGGAGAGAAGAAGCTCTTCAATGGCCAACAAGGAATTATTAT
TCAAAATTTTAGCACAAAGATCCATTCTCACTGTTACCAACGTGACACAGGAGCACTTCGGCA
ATTATACCTGTGTGGCTGCCAACAAGCTAGGCACAACCAATGCGAGCCTGCCTCTTAACCCT
CCAAGTACAGCCCAGTATGGAATTACCGGGAGCGCTGATGTTCTTTTCTCCTGCTGGTACCT
TGTGTTGACACTGTCTCTTTTACCAGCATATTCTACCTGAAGAATGCCATTCTACAATAAA
TTCAAAGACCCATAAAAGGCTTTTAAGGATTCTCTGAAAGTGCTGATGGCTGGATCCAATCT
GGTACAGTTTGTGTTAAAAGCAGCGTGGGATATAATCAGCAGTGCTTACATGGGGATGATCGCC
TTCTGTAGAATTGCTCATTATGTAAATACTTTAATTCTACTCTTTTTTTGATTAGCTACATTA
CCTTGTGAAGCAGTACACATTGTCCTTTTTTTAAGACGTGAAAGCTCTGAAATTACTTTTAG
AGGATATTAATTGTGATTTTCATGTTTGTAAATCTACAACTTTTCAAAGCATTTCAGTCATGGT
CTGCTAGGTTGCAGGCTGTAGTTTACAAAACGAATATTGCAGTGAATATGTGATTCTTTAA
GGCTGCAATACAAGCATTTCAGTTCCTGTTTTCAATAAGAGTCAATCCACATTTTACAAAGATG
CATTTTTTTCTTTTTTGATAAAAAAGCAAATAATATTGCCTTCAGATTATTTCTTCAAATA
TAACACATATCTAGATTTTTCTGCTTGCAATGATATTTCAGGTTTCAGGAATGAGCCTTGTAAT
ATAACTGGCTGTGCAGCTCTGCTTCTCTTTCTGTAAGTTTCAGCATGGGTGTGCCTTCATAC
AATAATATTTTTCTCTTTGTCTCCAATAATATAAAATGTTTTGCTAAATCTTACAATTTGA
AAGTAAAAATAAACAGAGTGATCAAGTTAAACCATACTATCTCTAAGTAACGAAGGAGC
TATTGGACTGTAAAAATCTCTTCTGCACTGACAATGGGGTTTGAGAATTTTGCCCCACACT
AACTCAGTTCTTGTGATGAGAGACAATTTAATAACAGTATAGTAAATATACCATATGATTTCT
TTTAGTTGTAGCTAAATGTTAGATCCACCGTGGGAAATCATTCCCTTTAAATGACAGCACA
GTCCACTCAAAGGATTGCCTAGCAATACAGCATCTTTTCTTTTCACTAGTCCAAGCCAAAAA
TTTTAAGATGATTTGTGAGAAAGGGCACAAAGTCCTATCACCTAATATTACAAGAGTTGGTA
AGCGCTCATCATTAATTTTATTTTGTGGCAGGTATTATGACAGTCGACCTGGAGGGTATGGA
TATGGATATGGACGTTCCAGAGACTATAATGGCAGAAACCAGGGTGGTTATGACCGCTACTC
AGGAGGAAATTACAGAGACAATTATGACAACCTGAAATGAGACATGCACATAATATAGATACA
CAAGGAATAATTTCTGATCCAGGATCGTCTTCCAAATGGCTGTATTTATAAAGGTTTTTGG
AGCTGCACTGAAGCATCTTATTTTATAGTATATCAACCTTTTGTTTTTTAAATTGACCTGCCA
AGGTAGCTGAAGACCTTTTAGACAGTTCATCTTTTTTTTTTAAATTTTTTCTGCCTATTTAA
AGACAAATTATGGGACGTTTGTCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 230

MMLLVQGACCSNQWLAAVLLSLCCLLPSCLPAGQSVDFPWAAVDNMMVRKGD TAVLR CYLED
GASKGAWLNRSSIIIFAGGDKWSVDPRVSISTLNKRDYSLQIQNVDVTDDGPYTCSVQTQHTP
RTMQVHLTVQVPPKIYDISNDMTVNEGTVNLTCLATGKPEPSISWRHISPSAKPFENGQYL
DIYGITRDQAGEYECSAENAVSFDPVRKVKV VVNFAPTIQEIKSGTVTPGRSGLIRCEGAGV
PPPAFEWYKGEKKLFNGQQGIIIQNFSTRSILTVTNVTQEHFGNYTCVAANKLGTTNASLPL
NPPSTAQYGITGSADVL FSCWYLVLTLS SFTSIFYLK NAILQ

Important features of the protein:

Signal peptide:

amino acids 1-31

Transmembrane domain:

amino acids 326-345

N-glycosylation sites.

amino acids 71-75, 153-157, 273-277, 284-288, 292-296, 305-309

Casein kinase II phosphorylation site.

amino acids 147-151, 208-212, 224-228

Tyrosine kinase phosphorylation site.

amino acids 178-186

N-myristoylation sites.

amino acids 7-13, 63-70, 67-73, 151-157, 239-245, 291-297,
302-308, 319-325

Myelin P0 protein:

amino acids 92-121

FIGURE 231

AGTGGTTCGATGGGAAGGATCTTTCTCCAAGTGGTTCCTCTTGAGGGGAGCATTTCTGCTGG
CTCCAGGACTTTGGCCATCTATAAAGCTTGGCAATGAGAAATAAGAAAATTCTCAAGGAGGA
CGAGCTCTTGAGTGAGACCCAACAAGCTGCTTTTCACCAAATTGCAATGGAGCCTTTCGAAA
TCAATGTTCCAAAGCCCAAGAGGAGAAATGGGGTGAACCTTCTCCCTAGCTGTGGTGGTCATC
TACCTGATCCTGCTCACCGCTGGCGCTGGGCTGCTGGTGGTCCAAGTTCTGAATCTGCAGGC
GCGGCTCCGGGTCCTGGAGATGTATTTCTCAATGACACTCTGGCGGCTGAGGACAGCCCGT
CCTTCTCCTTGCTGCAGTCAGCACACCCTGGAGAACACCTGGCTCAGGGTGCATCGAGGCTG
CAAGTCTGCAGGCCCCAACTCACCTGGGTCCGCGTCAGCCATGAGCACTTGCTGCAGCGGGT
AGACAACTTCACTCAGAACCCAGGGATGTTTCAAGATCAAAGGTGAACAAGGCGCCCCAGGTC
TTCAAGGTCAACAAGGGGGCCATGGGCATGCCTGGTGCCCCCTGGCCCCGCGGGACCACCTGCT
GAGAAGGGAGCCAAGGGGGCTATGGGACGAGATGGAGCAACAGGCCCCCTCGGGACCCCCAAGG
CCCACCGGGAGTCAAGGGAGAGGCGGGCCCTCCAAGGACCCCAGGGTGCTCCAGGGAAGCAAG
GAGCCACTGGCACCCCAGGACCCAAGGAGAGAAGGGCAGCAAAGGCGATGGGGGTCTCATT
GGCCCAAAGGGGAAACTGGAACCTAAGGGAGAGAAAGGAGACCTGGGTCTCCAGGAAGCAA
AGGGGACAGGGGCATGAAAGGAGATGCAGGGGTCTGAGGGCCTCCTGGAGCCCAGGGGAGTA
AAGGTGACTTCGGGAGGCCAGGCCCACCAGGTTTGGCTGGTTTTCTTGAGCTAAAGGAGAT
CAAGGACAACCTGGACTGCAGGGTGTTCCGGGCCCTCCTGGTGCAGTGGGACACCAGGTGC
CAAGGGTGAGCCTGGCAGTGCTGGCTCCCCCTGGGCGAGCAGGACTTCCAGGGAGCCCCGGGA
GTCCAGGAGCCACAGGCCTGAAAGGAAGCAAAGGGGACACAGGACTTCAAGGACAGCAAGGA
AGAAAAGGAGAATCAGGAGTTCCAGGCCCTGCAGGTGTGAAGGGAGAACAGGGGAGCCCAGG
GCTGGCAGGTCCCAAGGGAGCCCCCTGGACAAGCTGGCCAGAAGGGAGACCAGGGAGTGAAAG
GATCTTCTGGGGAGCAAGGAGTAAAGGGAGAAAAAGGTGAAAGAGGTGAAAACCTCAGTGTCC
GTCAGGATTGTCGGCAGTAGTAACCGAGGCCGGGCTGAAGTTTACTACAGTGGTACCTGGGG
GACAATTTGCGATGACGAGTGGCAAAATCTGATGCCATTGTCTTCTGCCGCATGCTGGGTT
ACTCCAAAGGAAGGGCCCTGTACAAAGTGGGAGCTGGCACTGGGCAGATCTGGCTGGATAAT
GTTCAGTGTCGGGGCACGGAGAGTACCCTGTGGAGCTGCACCAAGAATAGCTGGGGCCATCA
TGACTGCAGCCACGAGGAGGACGCAGGCGTGGAGTGCAGCGTCTGACCCCGGAAACCTTTCA
CTTCTCTGCTCCCGAGGTGTCCTCGGGCTCATATGTGGGAAGGCAGAGGATCTCTGAGGAGT
TCCCTGGGGACAACCTGAGCAGCCTCTGGAGAGGGGCCATTAATAAAGCTCAACATCATTGA

FIGURE 232

></usr/seqdb2/sst/DNA/Dnaseqs.full/ss.DNA68886

><subunit 1 of 1, 520 aa, 1 stop

><MW: 52658, pI: 9.16, NX(S/T): 3

MRNKKILKEDELLSETQQAAFHQIAMEPF EINVPKPKRRNGVNFSLAVVVIYLILLTAGAGL
LVVQVLNLQARLRVLEMYFLNDTLAAEDSPSFSLLQSAHPGEHLAQGASRLQVLQAQLTWVR
VSHEHLLQRVDNFTQNPGMFRIKGEQGAPGLQGHKGAMGMPGAPGPPGPPAEKGAKGAMGRD
GATGPSGPQGPPGVKGEAGLQGPQGAPGKQGATGTPGPQGEKSGKGDGGLIGPKGETGTKGE
KGD LGLPGSKGDRGMKGDAGVMGPPGAQGSKGDFGRPGPPGLAGFPGAKGDQGPGLQGVPG
PPGAVGHPGAKGEPGSAGSPGRAGLPGSPGSPGATGLKGSKGD TGLQGGQGRKGESGVPGPA
GVKGEQGS PGLAGPKGAPGQAGQKGDQGVKGSSGEQGVKGEKGERGENSVSVRIVGSSNRGR
AEVYYSGTWGTICDDEWQNSDAIVFCRMLGYSKGRALYKVGAGTGQIWLDNVQCRGTESTLW
SCTKNSWGHDCSHEEDAGVECSV

Transmembrane domain:

amino acids 47-66 (type II)

N-glycosylation sites.

amino acids 43-47, 83-87, 136-140

Tyrosine kinase phosphorylation site.

amino acids 432-440

N-myristoylation sites.

amino acids 41-47, 178-184, 253-259, 274-280, 340-346, 346-352,
400-406, 441-447, 475-481, 490-496, 515-521

Amidation site.

amino acids 360-364

Leucine zipper pattern.

amino acids 56-78

Speract receptor repeat

amino acids 422-471, 488-519

Clq domain proteins.

amino acids 151-184, 301-334, 316-349

Variable	Mean	SD	Min	Max
Age	30.5	4.2	18	45
Gender	0.5	0.5	0	1
Marital status	0.2	0.4	0	1
Education	12.5	1.5	9	16
Income	1500	500	500	3000
Health status	0.8	0.2	0	1
Smoking status	0.1	0.3	0	1
Alcohol consumption	0.2	0.4	0	1
Exercise frequency	0.5	0.5	0	1
Stress level	3.5	1.5	1	5
Sleep quality	4.0	1.0	2	5
Work satisfaction	3.8	1.2	1	5
Life satisfaction	4.2	1.0	2	5
Overall health	4.5	0.8	3	5
Physical health	4.3	0.7	3	5
Mental health	4.1	0.9	3	5
Social health	4.4	0.6	3	5
Emotional health	4.2	0.8	3	5
Behavioral health	4.0	0.9	3	5
Environmental health	4.3	0.7	3	5
Healthcare access	4.1	0.8	3	5
Health insurance	4.2	0.7	3	5
Healthcare quality	4.0	0.9	3	5
Healthcare cost	4.3	0.6	3	5
Healthcare satisfaction	4.1	0.8	3	5
Healthcare accessibility	4.2	0.7	3	5
Healthcare effectiveness	4.0	0.9	3	5
Healthcare safety	4.3	0.6	3	5
Healthcare equity	4.1	0.8	3	5
Healthcare transparency	4.2	0.7	3	5
Healthcare accountability	4.0	0.9	3	5
Healthcare innovation	4.3	0.6	3	5
Healthcare sustainability	4.1	0.8	3	5
Healthcare resilience	4.2	0.7	3	5
Healthcare adaptability	4.0	0.9	3	5
Healthcare inclusivity	4.3	0.6	3	5
Healthcare diversity	4.1	0.8	3	5
Healthcare collaboration	4.2	0.7	3	5
Healthcare partnership	4.0	0.9	3	5
Healthcare integration	4.3	0.6	3	5
Healthcare coordination	4.1	0.8	3	5
Healthcare communication	4.2	0.7	3	5
Healthcare education	4.0	0.9	3	5
Healthcare research	4.3	0.6	3	5
Healthcare development	4.1	0.8	3	5
Healthcare improvement	4.2	0.7	3	5
Healthcare optimization	4.0	0.9	3	5
Healthcare enhancement	4.3	0.6	3	5
Healthcare innovation	4.1	0.8	3	5
Healthcare sustainability	4.2	0.7	3	5
Healthcare resilience	4.0	0.9	3	5
Healthcare adaptability	4.3	0.6	3	5
Healthcare inclusivity	4.1	0.8	3	5
Healthcare diversity	4.2	0.7	3	5
Healthcare collaboration	4.0	0.9	3	5
Healthcare partnership	4.3	0.6	3	5
Healthcare integration	4.1	0.8	3	5
Healthcare coordination	4.2	0.7	3	5
Healthcare communication	4.0	0.9	3	5
Healthcare education	4.3	0.6	3	5
Healthcare research	4.1	0.8	3	5
Healthcare development	4.2	0.7	3	5
Healthcare improvement	4.0	0.9	3	5
Healthcare optimization	4.3	0.6	3	5
Healthcare enhancement	4.1	0.8	3	5
Healthcare innovation	4.2	0.7	3	5
Healthcare sustainability	4.0	0.9	3	5
Healthcare resilience	4.3	0.6	3	5
Healthcare adaptability	4.1	0.8	3	5
Healthcare inclusivity	4.2	0.7	3	5
Healthcare diversity	4.0	0.9	3	5
Healthcare collaboration	4.3	0.6	3	5
Healthcare partnership	4.1	0.8	3	5
Healthcare integration	4.2	0.7	3	5
Healthcare coordination	4.0	0.9	3	5
Healthcare communication	4.3	0.6	3	5
Healthcare education	4.1	0.8	3	5
Healthcare research	4.2	0.7	3	5
Healthcare development	4.0	0.9	3	5
Healthcare improvement	4.3	0.6	3	5
Healthcare optimization	4.1	0.8	3	5
Healthcare enhancement	4.2	0.7	3	5
Healthcare innovation	4.0	0.9	3	5
Healthcare sustainability	4.3	0.6	3	5
Healthcare resilience				

Variable	Mean	SD	Min	Max
Age	30.5	4.2	18	45
Gender	0.5	0.5	0	1
Marital status	0.2	0.4	0	1
Education	12.5	1.5	9	16
Income	1500	500	500	3000
Health status	0.8	0.2	0	1
Smoking status	0.1	0.3	0	1
Alcohol consumption	0.2	0.4	0	1
Exercise frequency	0.5	0.5	0	1
Stress level	3.5	1.5	1	5
Sleep quality	4.0	1.0	2	5
Work satisfaction	3.0	1.5	1	5
Life satisfaction	4.5	1.0	2	5
Depression score	1.5	1.0	0	3
Anxiety score	1.0	0.8	0	2
Resilience score	2.5	1.0	1	4
Optimism score	3.0	1.0	1	4
Gratitude score	3.5	1.0	1	4
Forgiveness score	3.0	1.0	1	4
Compassion score	3.0	1.0	1	4
Kindness score	3.0	1.0	1	4
Generosity score	3.0	1.0	1	4
Patience score	3.0	1.0	1	4
Self-control score	3.0	1.0	1	4
Emotional stability	3.0	1.0	1	4
Interpersonal skills	3.0	1.0	1	4
Problem-solving skills	3.0	1.0	1	4
Decision-making skills	3.0	1.0	1	4
Communication skills	3.0	1.0	1	4
Leadership skills	3.0	1.0	1	4
Teamwork skills	3.0	1.0	1	4
Conflict resolution skills	3.0	1.0	1	4
Time management skills	3.0	1.0	1	4
Organization skills	3.0	1.0	1	4
Planning skills	3.0	1.0	1	4
Goal setting skills	3.0	1.0	1	4
Self-motivation	3.0	1.0	1	4
Perseverance	3.0	1.0	1	4
Resilience	3.0	1.0	1	4
Optimism	3.0	1.0	1	4
Gratitude	3.0	1.0	1	4
Forgiveness	3.0	1.0	1	4
Compassion	3.0	1.0	1	4
Kindness	3.0	1.0	1	4
Generosity	3.0	1.0	1	4
Patience	3.0	1.0	1	4
Self-control	3.0	1.0	1	4
Emotional stability	3.0	1.0	1	4
Interpersonal skills	3.0	1.0	1	4
Problem-solving skills	3.0	1.0	1	4
Decision-making skills	3.0	1.0	1	4
Communication skills	3.0	1.0	1	4
Leadership skills	3.0	1.0	1	4
Teamwork skills	3.0	1.0	1	4
Conflict resolution skills	3.0	1.0	1	4
Time management skills	3.0	1.0	1	4
Organization skills	3.0	1.0	1	4
Planning skills	3.0	1.0	1	4
Goal setting skills	3.0	1.0	1	4
Self-motivation	3.0	1.0	1	4
Perseverance	3.0	1.0	1	4
Resilience	3.0	1.0	1	4
Optimism	3.0	1.0	1	4
Gratitude	3.0	1.0	1	4
Forgiveness	3.0	1.0	1	4
Compassion	3.0	1.0	1	4
Kindness	3.0	1.0	1	4
Generosity	3.0	1.0	1	4
Patience	3.0	1.0	1	4
Self-control	3.0	1.0	1	4
Emotional stability	3.0	1.0	1	4
Interpersonal skills	3.0	1.0	1	4
Problem-solving skills	3.0	1.0	1	4
Decision-making skills	3.0	1.0	1	4
Communication skills	3.0	1.0	1	4
Leadership skills	3.0	1.0	1	4
Teamwork skills	3.0	1.0	1	4
Conflict resolution skills	3.0	1.0	1	4
Time management skills	3.0	1.0	1	4
Organization skills	3.0	1.0	1	4
Planning skills	3.0	1.0	1	4
Goal setting skills	3.0	1.0	1	4
Self-motivation	3.0	1.0	1	4
Perseverance	3.0	1.0	1	4
Resilience	3.0			

[illegible]

<subunit 1 of 1, 98 aa, 1 stop

MKLMVLVFTIGLTLLLG VQAMPANRLSCYRKILKDHNCHNLPEGVADLTQIDVNVQDHFWDG

Important features:

amino acids 1-20

amino acids 72-76

amino acids 63-71

FIGURE 235

CCCACGCGTCCGCGGACGCGTGGGCTGGACCCAGGTCTGGAGCGAATTCAGCCTGCAGGG
CTGATAAGCGAGGCATTAGTGAGATTGAGAGAGACTTTACCCCGCCGTGGTGGTTGGAGGGC
GCGCAGTAGAGCAGCAGCACAGGCGCGGGTCCCGGGAGGCCGGCTCTGCTCGCGCCGAGATG
TGAATCTCCTTCACGAAACCGACTCGGCTGTGGCCACCGCGCGCCCGCCGCGCTGGCTGTG
CGCTGGGGCGCTGGTGTCTGGCGGGTGGCTTCTTTCTCCTCGGCTTCTCTTCGGGTGGTTTA
TAAAATCCTCCAATGAAGCTACTAACATTACTCCAAAGCATAATATGAAAGCATTTTTGGAT
GAATTGAAAGCTGAGAACATCAAGAAGTTCTTACATAATTTTACACAGATACCACATTTAGC
AGGAACAGAACAAAACCTTTCAGCTTGCAAAGCAAATTCATCCAGTGGAAGAATTTGGCC
TGGATTCTGTTGAGCTAGCTCATTATGATGTCTGTTGTCCTACCCAAATAAGACTCATCCC
AACTACATCTCAATAATTAATGAAGATGGAAATGAGATTTTCAACACATCATTATTTGAACC
ACCTCCTCCAGGATATGAAAATGTTTTCGGATATTGTACCACCTTTCAGTGCTTCTCTCCTC
AAGGAATGCCAGAGGGCGATCTAGTGTATGTTAACTATGCACGAACTGAAGACTTCTTTAA
TTGGAACGGGACATGAAAATCAATTGCTCTGGGAAAATTGTAATTGCCAGATATGGGAAAGT
TTTCAGAGGAAATAAGGTTAAAAATGCCAGCTGGCAGGGGCCAAAGGAGTCATTCTCTACT
CCGACCTTGCTGACTACTTTGCTCCTGGGGTGAAGTCCTATCCAGACGGTTGGAATCTTCCT
GGAGGTGGTGTCCAGCGTGGAATATCCTAAATCTGAATGGTGCAGGAGACCTCTCACACC
AGGTTACCCAGCAAATGAATATGCTTATAGGCGTGGAATTGCAGAGGCTGTTGGTCTTCCAA
GTATTCCTGTTTCATCCAATTGGATACTATGATGCACAGAAGCTCCTAGAAAAAATGGGTGGC
TCAGCACCACCAGATAGCAGCTGGAGAGGAAGTCTCAAAGTGCCCTACAATGTTGGACCTGG
CTTTACTGGAACTTTTTCTACACAAAAGTCAAGATGCACATCCACTCTACCAATGAAGTGA
CGAGAATTTACAATGTGATAGGTACTCTCAGAGGAGCAGTGGAAACCAGACAGATATGTCATT
CTGGGAGGTACCCGGGACTCATGGGTGTTTGGTGGTATTGACCCTCAGAGTGGAGCAGCTGT
TGTTTCATGAAATTGTGAGGAGCTTTGGAACACTGAAAAAGGAAGGTGGAGACCTAGAAGAA
CAATTTTGTGTTGCAAGCTGGGATGCAGAAGAATTTGGTCTTCTTGGTTCTACTGAGTGGGCA
GAGGAGAATTCAGACTCCTTCAAGAGCGTGGCGTGGCTTATATTAATGCTGACTCATCTAT
AGAAGGAAACTACACTCTGAGAGTTGATTGTACACCGCTGATGTACAGCTTGGTACACAACC
TAACAAAAGAGCTGAAAAGCCCTGATGAAGGCTTTGAAGGCAAATCTCTTTATGAAAGTTGG
ACTAAAAAAGTCCTTCCCCAGAGTTCAGTGGCATGCCAGGATAAGCAAATTGGGATCTGG
AAATGATTTTGAAGTGTCTTCCAACGACTTGGAAATTGCTTCAGGCAGAGCACGGTATACTA
AAAATTGGGAAACAAACAAATTCAGCGGCTATCCACTGTATCACAGTGTCTATGAAACATAT
GAGTTGGTGGAAAAGTTTTATGATCCAATGTTTAAATATCACCTCACTGTGGCCCAGGTTG
AGGAGGGATGGTGTGTTGAGCTAGCCAATTCATAGTGCTCCCTTTTGATTGTGAGATTATG
CTGTAGTTTTAAGAAAGTATGCTGACAAAATCTACAGTATTTCTATGAAACATCCACAGGAA
ATGAAGACATACAGTGTATCATTTGATTCACTTTTTTCTGCAGTAAAGAATTTTACAGAAAT
TGCTTCCAAGTTCAGTGAGAGACTCCAGGACTTTGACAAAAGCAACCAATAGTATTAAGAA
TGATGAATGATCAACTCATGTTTCTGGAAAGAGCATTATTTGATCCATTAGGGTTACCAGAC
AGGCCTTTTTTATAGGCATGTCATCTATGCTCCAAGCAGCCACAACAAGTATGCAGGGGAGTC
ATTCCCAGGAATTTATGATGCTCTGTTTGTATTTGAAAGCAAAGTGGACCCTTCCAAGGCCT
GGGGAGAAGTGAAGAGACAGATTTATGTTGCAGCCTTCACAGTGCAGGCAGCTGCAGAGACT
TTGAGTGAAGTAGCCTAAAGAGGATTTTTTATAGAGAATCCGTATTGAATTTGTGTGGTATGTCA
CTCAGAAAGAATCGTAATGGGTATATTGATAAATTTTAAAATTGGTATATTTGAAATAAAGT
TGAATATTATATATAA

FIGURE 236

></usr/seqdb2/sst/DNA/Dnaseqs.full/ss.DNA52756

><subunit 1 of 1, 750 aa, 1 stop

><MW: 84305, pI: 6.93, NX(S/T): 10

MWNLLHETDSAVATARRPRWLCAGALVLAGGFLLGFLFGWFIKSSNEATNITPKHNMKAFL
DELKAENIKKFLHNFTQIPHLAGTEQNFQLAKQIQSQWKEFGLDSVELAHYDVLLSYPNKTH
PNYISIINEDGNEIFNTSLFEP PPPGYENVSDIVPPFAFSPQGMPEGDLVYVNYARTEDFF
KLERDMKINCSGKIVIRYGVFRGNKVNAQLAGAKGVILYSDPADYFAPGVKSYPDGWNL
PGGGVQRGNILNLNGAGDPLTPGYPANAYRRGIAEAVGLPSIPVHPIGYYDAQKLLLEKMG
GSAPPDSSWRGSLKVPYNVGPFTGNFSTQKVKMHIHSTNEVTRIYNVIGTLRGAVEPDYV
ILGGHRDSWVFGGIDPQSGAAVVHEIVRSFGTLKKEGWRPRRTILFASWDAEEFGLLGSTEW
AEENSRLQLQERGVAYINADSSIEGNYTLRVDCTPLMYSLVHNLTKELKSPDEGFEGKSLYES
WTKKSPSPEFSGMPRISKLGSGNDFEVFFQRLGIASGRARYTKNWETNKFSGYPLYHSVYET
YELVEKFYDPMFKYHLTVAQVRGGMVFELANSIVLPFDCRDYAVVLRKYADKIYSISMKHPQ
EMKTYSVSFDLSLFSVKNFTEIASKFSERLQDFDKSNPIVLRMMNDQLMFLERAFIDPLGLP
DRPFYRHVIYAPSSHNKYAGESFPGIYDALFDIESKVDPSKAWGEVQRQIYVAAFTVQAAAE
TLSEVA

Signal sequence:

amino acids 1-40

N-glycosylation sites.

amino acids 76-80, 121-125, 140-144, 153-157, 195-199, 336-340,
459-463, 476-480, 638-642

Tyrosine kinase phosphorylation sites.

amino acids 363-372, 605-613, 606-613, 617-626

N-myristoylation sites.

amino acids 85-91, 168-174, 252-258, 256-262, 282-288, 335-341,
360-366, 427-433, 529-535, 707-713